DOCUMENT RESUME

ED 134 193 IR 004 370

AUTHOR	McCallum, Ian S.
TITLE	The On-Line ERIC Project at the National Library of Australia.
INSTITUTION	National Library of Australia, Canberra.
PUB DATE	76
NOTE	132p.
EDRS PRICE	MF-$0.83 HC-$7.35 Plus Postage.
DESCRIPTORS	Data Bases; *Educational Resources; *Information Retrieval; *On Line Systems; Questionnaires; Telecommunication
IDENTIFIERS	*Australia; *Educational Resources Information Center; ERIC

ABSTRACT

This report describes the On-Line ERIC Project conducted jointly by the National Library of Australia and by IBM Australia Ltd. between March and May 1976. The Project involved a telecommunications network with video terminals and printers linking the National Library, the Royal Melbourne Institute of Technology Central Library, and Macquarie University Library to an IBM computer in Canberra. Citations from two and a half years of the ERIC data base were available through the Storage and Information Retrieval System (STAIRS) software package for on-line searching for three hours per day for three days each week for three months. Users from various backgrounds responded favorably to the system's interactive searching capability as revealed in the survey results. Retrieval costs in manual, batch, and on-line processing mode were compared. An instructional package of how to use STAIRS is included in the appendix. (Author/SC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED134193

IR

THE ON-LINE ERIC PROJECT AT THE NATIONAL LIBRARY OF AUSTRALIA

An experiment to extend the investigations of the ERIC Research Project through an on-line education information service using the STAIRS processing system to interrogate the ERIC data base

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

Ian S. McCallum

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Ian S. McCallum
Education Information Service
National Library of Australia

Canberra, National Library of Australia, 1976

IR004370

National Library of Australia. Education Information Service.
The On-line ERIC Project at the National Library of Australia: an experiment to extend the investigations of the ERIC Research Project through an on-line education information service using the STAIRS processing system to interrogate the ERIC data base/ [by] Ian S. McCallum.- Canberra: National Library of Australia, 1976.-
2 cards.
Microfiche.
ISBN 0 642 99078 6.

1. ERIC.
2. Education - Information services - Australia.
3. Information storage and retrieval systems - Education. I. McCallum, Ian Stanley, 1947-.
II. Title.
029.937

ACKNOWLEDGEMENTS

Like most projects, the On-line ERIC Project was a team effort. The members of the team were: Peter Urquhart from IBM Canberra; Ray Penn at Macquarie University; Ann Wilson at the Royal Melbourne Institute of Technology; and Sandra Duffield, Diana Killen and Lena Smith from the National Library's Education Information Service. Sandra drafted parts A and B of section III of this report. The success of the Project was attributable to the abilities of these people.

I.S. McC.

ABSTRACT

This report describes the On-line ERIC Project conducted jointly by the National Library of Australia and by IBM Australia Ltd between March and May 1976. The Project involved a telecommunications network with video terminals and printers linking the National Library, the Royal Melbourne Institute of Technology Central Library, and Macquarie University Library to an IBM computer in Canberra. Citations from two and a half years of the ERIC data base were available through the STAIRS software package for on-line searching for three hours per day for three days each week for three months.

Almost without exception, users from universities, colleges, government departments, libraries, industrial, commercial and professional associations placed an extremely high value on the system's interactive searching capability. There was no doubt that the Project represented a significant advance in the development of improved information services for educators.

TABLE OF CONTENTS

I.	Setting up the Project	6
II.	On-line ERIC in action - the network	18
III.	On-line ERIC in action - the computing system	22
IV.	User response	46
V.	Conclusions	85
VI.	Appendix 1: Easy steps to STAIRS	88

I. SETTING UP THE PROJECT

The On-line ERIC Project started in March 1975 with a suggestion made at a National Library senior staff meeting. At that time the Library was sponsoring several ALBIS (Australian Library Based Information System) feasibility studies ranging from an investigation of tariff policies to the development of plans for the introduction of subject-oriented resource sharing networks. The attempt at co-operative formulation of a national plan for the establishment of up to 40 separate subject networks reflected the National Library's ability to stimulate the development of Australian library systems, yet the National Library Council held few illusions about the magnitude of the task. The Library's Director-General, Dr George Chandler, supported by the Council and the Minister, was determined that not all ALBIS projects should be visionary and long-term, and he directed the ALBIS study teams to pay particular attention to the identification of short-term practical advantages and reminded those involved that in order to keep faith with the library community, practical suggestions and visible output were highly desirable. This pragmatic out-look characterised the Library staff members who had been seconded to ALBIS.

When On-line ERIC was first mooted, it received a sympathetic hearing: the Library's Education Information Service (EIS) offered among other things a successful current awareness service from the ERIC data base but had failed to attract an additional $8000 p.a. to cover data processing costs for the addition of a retrospective capability. The On-line ERIC Project was therefore envisaged as a way of meeting existing

user pressure for retrospective searches and also as the vehicle for an investigation into the practicability of on-line information services in education. For some time the EIS staff had been concerned with the apparent linear relationship between the number of active profiles and the number of staff required to service those profiles (roughly one staff member for each 100-150 profiles), and they felt that while their decentralised profile writing seminars were at least an attempt at spreading the cost of information delivery, this procedure was really a stop-gap measure. Intuitively it seemed that the ideal solution embodying unrestricted access to the data base as well as accurate and effective profile writing depended upon training the users themselves to operate remotely-located video or hard copy terminals connected via data transmission lines to a single computer. It was felt that such a network represented the best solution to the long-term problem of providing an education information service that had the capacity to process the requests it received.

For the Director-General as the executive head of ALBIS, the On-line ERIC Project met the criteria of a useful investigation: there was plenty of evidence of user demand for retrospective access to the ERIC data base; STAIRS (Storage and Information Retrieval System), the IBM software package proposed for the Project, would accept as input for building the data base magnetic tapes which had already been produced for normal SDI processing with TEXTPAC - another IBM system; and the significance of remotely-located user-operated terminals was obvious for the elaboration of an education information network.

In addition there existed the possibility that IBM's Canberra-based Systems Development Institute which had supported the data processing associated with the ERIC Research Project* which preceded the establishment of the EIS, might be interested in providing similar assistance for the On-line Project.

The original proposal was framed within the context of ALBIS: it mentioned the collection of data as input to the ALBIS studies of potential information networks and suggested the addition of an on-line searching facility to the EIS. The proposer felt that such a project would provide an opportunity to assess the effectiveness of interactive searching and would yield valuable data on the operation of network services. Accordingly, a three-node network was proposed. Video terminals in Melbourne, Sydney, and Canberra would be linked by high-speed data lines to a computer in Canberra where the latest two years of the ERIC data base would be available for searching for nine hours a week for three months.

In late March 1975, Colin Freeman, Bob Simmons and Ian McCallum from the National Library met with Gordon Undy from IBM to decide whether or not the Project was feasible. The meeting decided that the proposal warranted closer attention and made a preliminary estimate of the staff and data processing resources required. The first cost estimate was as follows:

*McCallum, Ian S. and W. Neville Holmes. The ERIC Research Project at the National Library of Australia: an experiment in providing information services from the ERIC data base using the TEXTPAC processing system. Canberra, National Library of Australia, 1975. 53p. One microfiche.

(i)	rental of 4800 bps lines for 4 months	13,000
(ii)	rental of terminal equipment for 4 months	12,000
(iii)	rental of direct access storage for 4 months	3,000
(iv)	data processing for 4 months	8,000
(v)	connect-time for 4 months	4,000
	TOTAL:	$40,000

Since the Director-General had indicated that he would only be prepared to allocate up to $25,000 for this particular ALBIS project, Gordon Undy undertook to determine his company's attitude to supporting another joint venture between the Library and IBM's Systems Development Institute. In an earlier submission to the Library on ALBIS, IBM had intimated that they would consider favourably joint projects of the type envisaged in the proposal, and this generous attitude encouraged the Library to continue planning. A more detailed proposal was then submitted to Dr Chandler. This quotation is from the original document:

> "It is proposed that two years of the ERIC data base be stored on-line at the IBM Systems Development Institute in Canberra and be accessible in three 3-hour slots per week for three months. On-line access would be provided in Canberra, Sydney and Melbourne at the National Library and possibly at Macquarie University Library and the Royal Melbourne Institute of Technology."

At this early stage it was felt that the simplest and therefore cheapest network to establish would link Sydney, Canberra and Melbourne. To have extended the telecommunications lines to Brisbane and Adelaide would have pushed the costs beyond ALBIS' ability to fund the Project. As it was, the Project only became a reality through IBM's interest.

Macquarie University with its widely-respected education faculty and its comprehensive collection of ERIC mirofiche was thought to be the logical site for the Sydney terminal. In Melbourne, La Trobe University had a growing reputation as a centre for educational studies, Monash University had a full set of ERIC microfiche and a known interest in as well as some success with library data processing, and Melbourne University had probably the best collection of educational material. In the end, the Royal Melbourne Institute of Technology (RMIT), was chosen for its central location, its interests in and experience with educational technology, its capable staff, and also becuase it represented an organisation different from a university.

The proposal then set out the details of a cooperative project. It was envisaged that the National Library, through ALBIS, would pay $25,000 for line and equipment rentals and IBM would carry about $15,000 in data processing costs. The original $40,000 estimate was retained. This was approved in principle by the Director-General, subject to a later report.

In mid-April he wrote to the Systems Development Institute, seeking IBM's support for a three month joint study beginning in August 1975, and raised the question of the availability of suitable terminal equipment - IBM 3275 visual displays and IBM 3284 matrix printers. The Director-General also asked the Director-General of Posts and Telegraphs, Mr. E.F. Lane, whether the then Australian Post Office could assist the proposed study into the information use patterns of on-line network users by installing the necessary telecommunications network. Like the letter to IBM, this letter was written within the context of the ALBIS feasibility studies, seeking assistance that had already been offered.

IBM then advised the Library that a definite delivery date for terminal equipment was not possible until the Library placed a firm order for temporary rental of the screens and printers. On 21 April the Library placed an order for temporary rental of three visual display units and three matrix printers for a monthly charge of $1371.60. In April the Australian Post Office advised the Library that they could install the necessary lines and modems in time for an August start to the Project.

Then in mid-May IBM informed the Library that the terminal equipment would not be available until December-January. The Library's Corporate Board of Management (the Director-General in consultation with the Directors of each branch of the Library) considered this new development and decided to continue with the Project, accepting the delayed start. At the same time IBM officially agreed to support the study:

"We would find this a very interesting extension of the ERIC project we carried out with you and we would like to proceed with you under similar terms to the previous project." (Letter from IBM to the Director-General, 18 May 1975).

Three months after the intial suggestion the Project was taking shape. Bill Thorn, the Director of the Australian National Social Sciences Library, a branch of the National Library, then wrote to Eoin Wilkinson, Macquarie University Librarian, and to Jack Ward, RMIT librarian, to outline the Project and to propose that the Sydney and Melbourne terminals be installed at Macquarie and at RMIT. Both Wilkinson and Ward were enthusiastic, and readily accepted the offer to take part in this ALBIS study. The Australian Post Office (now Telecom Australia) was advised of the postponed commencement date and IBM guaranteed a start date of the first week of February 1976.

The Telecom Australia quotation for lines and modems for four months was $6986 plus installation and recovery. In consultation with IBM's computing centre manager, it was agreed that the Project would be operated from 8am to 11 am on Tuesdays, Wednesdays and Thursdays for March, April and May with February for test and development work. IBM assigned one of their Systems Engineers, Peter Urquhart, to the Project and he began to build the STAIRS data base from output tapes from EIS SDI production runs. By September Telecom had begun installing the telecommunications lines and modems.

The staff of the EIS turned their attention to the development of training materials for on-line users. Following extensive literature searches, including some correspondence with the State University of New York (who were using STAIRS to search the ERIC and MEDLARS data bases), it was decided that the single most effective user aid would be a simple but comprehensive manual. The EIS also started work on a tape/slide sequence designed to depict the interaction between system and user as a little less than forbidding.

At the same time Peter Urquhart was examining the possibility of storing the Thesaurus of ERIC Descriptors on-line as an aid to interactive searching. A thesaurus tape was loaded into the system as a data base but it was obvious that without the facility to search and expand the reference structure for particular descriptors, a machine-readable on-line thesaurus offered few advantages over the printed version. Since the considerable programming effort required to transfer the thesaurus tape into an on-line searching aid was estimated to exceed the advantages accruing to such a short-term project, this activity was suspended. IBM then advised a minor variation in the times the system would be available to users: from 8a[illegible] 11am to a new time of 9.30 am to 12:30 pm. The slightly changed times conformed more closely to IBM operating system requirements and processing loads and caused no problems for either the Library or for Telecom.

At the National Library on the 8th of December 1975, a meeting between Ann Wilson, RMIT's Senior Reference Librarian, Margaret Cameron, Macquarie University's Reader Services Librarian, Peter Urquhart and EIS staff was held to discuss the state of the Project.

Drafts of the users' manual, the terminal users' questionnaire and a statement of objectives were considered. A publicity program was planned. It was felt that within the 80 hours of on-line time available to each terminal over the course of the Project any investigation of the information use patterns of terminal users should involve at least three separate sessions with a first session of 60 minutes, a second session of 30 minutes and subsequent sessions of 15 minutes. By the end of the third session it was expected that users would be reasonably proficient in operating the system and that their questionnaire responses at the completion of each session would yield the necessary feedback for a thorough evaluation. This meant that each location would have to plan for about 30 "serious users" whose full cooperation would be sought, and left roughly 20 hours terminal time for each centre to use as it saw fit - for demonstration and for its own purposes. Selection of users was to be left to each location but was to be partially based upon NLA-supplied lists of current ERIC users since IBM had requested some overlap with the user group for the earlier ERIC Research Project. It was also decided that the NLA should carry primary responsibility for the provision of document back-up.

The on-line data base would comprise Resources in Education (RIE) and Current Index to Journals in Education (CIJE) for 1974 and 1975, plus updating tapes for 1976 when received from the U.S. Project co-ordinators, to be responsible for user training and system operation, were appointed: Ray Penn in Sydney; Ann Wilson in Melbourne; and Ian McCallum in Canberra.

The agreed statement of objectives read as follows:

"THE ON-LINE ERIC PROJECT - STATEMENT OF OBJECTIVES

Between September 1972 and January 1974 the National Library of Australia and IBM Australia conducted a joint study into the feasibility of computer-based current awareness and retrospective search services. In 1974 the ERIC Research Project, as it was called, became the operational Education Information Service. Now, three years after the start of the Research Project, the National Library realises that to meet the demand for information services based on ERIC and other education data, new processing techniques must be carefully examined. IBM have agreed to support a further study, this time involving on-line information retrieval. Called the On-line ERIC Project, the joint venture provides for the latest two years of the ERIC files to be simultaneously accessible to users in Canberra, Melbourne and Sydney. Printers and cathode ray terminals at the National Library, the Royal Melbourne Institute of Technology and at Macquarie University will be available to the user group for three hours per day for three days per week for three months beginning in March 1976.

PRIMARY OBJECTIVES

1. To extend and consolidate the investigations of the ERIC Research Project by examining the applicability of on-line information retrieval techniques to the subject areas of education.

2. To investigate the extent to which an on-line system such as STAIRS is able to meet the data processing requirements of the Education Information Service for the following five years, and related to this objective, to investigate the extent to which an on-line system is an appropriate computing system for supporting education network development.

3. To analyse the costs and benefits of on-line education information retrieval compared with off-line education information retrieval, and to assess the results of such an analysis in terms of education network developments.

SUBSIDIARY OBJECTIVES

1. To examine the effect of an on-line information retrieval service on the information-gathering practices of its users.

2. To compare the relative effectivess of search formulations written directly by users with search formulations written by search analysts, and, in so doing, to assess the feasibility of a decentralised, remotely-accessed education information network.

3. To stimulate Australian interest in, and enhance Australian access to, the results of foreign educational research.

4. To assess the extent to which the ERIC data base meets the information needs of Australians working in the field of education."

The objectives addressed themselves to two major tasks:

(i) to examine the suitability of on-line information retrieval as the basis for development of a resource sharing network in education;

(ii) to assess the effectiveness of on-line information retrieval as a short term means of improving information delivery in education.

In January 1976 IBM installed the screens and printers in Canberra, Melbourne and Sydney. Copies of the users' manual Easy steps to STAIRS *, the tape/slide sequence STAIR cases , the pre- and post-session questionnaires, documentation on STAIRS and the terminal equipment, co-ordinators' manuals, and paper and ribbons for the printers were all despatched to Macquarie and RMIT over the next few weeks. Then in February users were invited to join the Project. By the end of that month all was in readiness for the start: equipment and lines had been set up and tested; the National Library had issued a press release which had been taken up by ABC radio in Canberra, by the Canberra Times, The Australian and The Age; and users had been scheduled for three months of terminal sessions.

*Attached as Appendix 1.

II. ON-LINE ERIC IN ACTION - THE NETWORK

The on-line computing network used for the Project was, by comparison with an off-line system, a very complicated device. The IBM computer in Canberra is used mainly for system development and its hardware and software are unusually complex, and in addition there were three 4800 bit per second lines with modems at both ends supplying computing power to video displays and printers at three locations. Consequently the probability of the network functioning in a less than optimum manner was high. Of the 347 hours of scheduled system availability (9 hours each week in three locations for three months), system failures accounted for almost 56 hours or 16% of lost time. The components of the network fared as follows:

LOCATION	SCHEDULED TIME (hours)	TIME AVAILABLE (hours)	%FAILURE
CANBERRA	118	105	11
MELBOURNE	113	97	14
SYDNEY	116	89	23

It is convenient to explain the shortfalls in performance in terms of three broad categories:

1. Faults in the telecommunications system. This category includes problems with modems, lines and terminal equipment.

2. Faults in the computing system, including both hardware and software problems. In this category fall the range of failures from breakdowns in the physical units of the IBM 370/158 computer in Canberra to problems with the operating system, with the communications control software and with STAIRS (Storage and Information Retrieval System) itself.

3. The third category is probably best described as "marginal conditions", and includes two major problem sources: operator error either at the terminal or in the machine-room; and poor response time. This last marginal condition, while not strictly a failure in the computing system did limit the time available to network users.

The table below gives details by location:

LOCATION	CATEGORY 1 TELE-COMMUNICATIONS	CATEGORY 2 COMPUTING SYSTEM	CATEGORY 3 MARGINAL CONDITIONS	TOTAL
CANBERRA	16	76	8	100%
MELBOURNE	35	59	6	100%
SYDNEY	63	30	7	100%

% of time lost through faults at each location

The figures in this table refer to the percentages of time lost to the Project through faults ascribed to the three categories. They are not related to the incidence of faults within each category. For example, while 59% of the time lost in Melbourne was caused by faults in the computing system, and 35% of the time lost was caused by faults in the telecommunications system, there were many more problems with lines and modems than with the computing system. So, in interpreting the table, it must be remembered that faults occurring in the computing system, although relatively infrequent compared with line and modem faults, took longer to rectify. In fact the facilities of CICS (Customer Information Control System), the software under which STAIRS ran on the computer, enabled the project team to quickly re-enable a defective line as soon as it went out of service, thereby minimising the interruption.

Of the 39 on-line sessions available to the Project over March, April and May only 6 sessions (15%) were trouble-free. The most persistent and frustrating problems occurred with the link from Canberra to Sydney. Despite the fact that at least five formal complaints were made to Telecom in Sydney over the three months of the Project, and despite Telecom's rewiring and replacement of modems, the line from Canberra to Sydney was never as dependable as the lines to the National Library and to RMIT.

This meant that whenever the Sydney line went out of service, users on the Canberra terminal were interrupted while an EIS staff member used the Canberra terminal as a master terminal to re-enable the Sydney line. And since the Sydney terminal or line went out of service more than 30 times over the 39 Project days, at an average of 5 minutes to re-enable the link each time, this activity restricted the time available to Canberra users. By comparison the Melbourne line went out of service only 15 times and the Canberra link had to be re-enabled on only two occasions.

III. ON-LINE ERIC IN ACTION - THE COMPUTING SYSTEM

Thus far this report has described how the Project originated and how it progressed. This section will describe how the computing system was used and how it was evaluated.

1. A typical terminal session

Let us assume that a user wants to retrieve material on the relationship between communication skills and egocentrism in young children. What follows is essentially a transcript of the dialogue between user and system ("S" indicates system response, "U" indicates user instruction, comments are in parentheses).

S: Query 1

(The system tells the user it is ready and prompts him/her for the topic.)

U: (COMMUNICATION ADJ SKILLS).DESCRIPTORS.

(The user enters his/her first search statement requesting retrieval of citations containing the phrase "communication skills". "ADJ" indicates that the words "communication" and "skills" should occur adjacent to each other and in the order specified. ".DESCRIPTORS." retricts the search to the descriptor or subject term field of the citation. This device was usually employed where the user had consulted the Thesaurus of

ERIC descriptors and wished to retrieve citations with terms the original ERIC indexer had considered sufficiently indicative of concepts within citations to warrant specifying those terms in the descriptor fields of the citations.)

S: Result

Number of occurrences = 292
Number of documents = 292

(The system responds with a tally of the number of times the phrase "communication skills" occurs within all the citations in the file being searched, and also with a tally of the number of citations in which the phrase appears.)

U: Query 2 (VERBAL ADJ ABILITY).DESCRIPTORS.

(The user enters his/her second query, also limited to the descriptor field.)

S: Result

Number of occurrences = 75
Number of documents = 75

U: Query 3 1 OR 2

(The user then decides to search for all citations indexed with the phrases "communication skills" or "verbal ability".)

S: Result

Number of occurrences = 367
Number of documents = 350

(This time there are fewer citations than occurrences and this conveys to the user the information that 13 citations contain both phrases.)

U: Query 4 EGOCENT$

(Having dealt with the communication skills and related verbal ability concepts, the user then searches on the "egocentric" aspect of his/her query. The "$" will mask subsequent letters in the word so that citations containing the words "egocentric" and "egocentrism" will be retrieved. Note that there is no limitation on which parts of the citation the system is instructed to search. If the word stem "egocent" is found anywhere (in any field) of a citation, the item will be retrieved.)

S: Result

Number of occurrences = 50
Number of documents = 29

U: Query 5 (SELF ADJ (CONCEPT OR ESTEEM)).DESCRIPTORS.

(The user then searches the descriptor field for the phrases "self concept" or "self esteem".)

S: Result

Number of occurrences = 716
Number of documents = 649

U: Query 6 4 OR 5

(The user decides to combine queries 4 and 5.)

S: Result

Number of occurrences = 766
Number of documents = 673

U: Query 7 3 AND 6

(The user combines all subsequent queries so that if a citation contains either of the terms "communication skills" or "verbal ability" as well as either of the terms "egocent$" or "self concept" or "self esteem", then the items which match the search strategy will be retrieved.)

S: Result

Number of occurrences = 11
Number of documents = 11

The user then has the options of displaying the citations one by one on the video screen, printing any considered relevant, or expanding the search strategy using terms from the text of citations. A search like this would take about 10 minutes with the system responding to each new instruction within a few seconds. Most of the 10 minutes

would be taken up by the user typing in his/her request. As we shall see, it was the speed of the system response which impressed users most.

2. Terminal session costs

The data processing cost of providing the on-line service to three centres simultaneously for three hours duration, ranged from $470 to $940. The variation is accounted for in terms of:

(i) the total number of users using the system within the three, three-hour time slots - this ranged from seven to 11

(ii) the way the system was used, ie. whether large numbers of citations were searched to produce large output files or whether small numbers of citations were searched to produce small output files

(iii) the number of citations users decided to print

The following tables illustrate two typical sessions.

Session 1. 6 April 1976

LOCATION	SESSION TIME (minutes)	CITATIONS		
		RETRIEVED	RELEVANT	PRINTED
CANBERRA (3 users)	110	163	156	163
MELBOURNE (4 users)	95	73	59	66
SYDNEY (2 users)	150	590	339	339

COMPUTING COSTS (at IBM service bureau rates)

On-line sessions:	$510
Printing results	$200

LINE CHARGES (at 1/3 of the weekly rental charge):

Canberra	$12
Melbourne	$65
Sydney	$60

EQUIPMENT RENTAL FOR THREE LOCATIONS

(at 1/12 monthly rate): $130

TOTAL SESSION COSTS (excluding staff salaries and administrative overhead): $977

COST PER RETRIEVED CITATION: $1.18

Session 2. 4 May 1976

LOCATION	SESSION TIME (minutes)	CITATIONS		
		RETRIEVED	RELEVANT	PRINTED
CANBERRA (4 users)	170	87	80	87
MELBOURNE (2 users)	65	130	80	80
SYDNEY (2 users)	130	168	121	123

COMPUTING COSTS (at IBM service bureau rates):

On-line sessions	$400
Printing results	$125

LINE CHARGES (1/3 weekly rate):

Canberra	$12
Melbourne	$65
Sydney	$60

EQUIPMENT RENTAL FOR THREE LOCATIONS (at 1/12 monthly rate): $130

TOTAL SESSION COSTS (excluding staff salaries and administrative overhead): $792

COST PER RETRIEVED CITATION: $2.06

3. STAIRS and the supporting computer system

RIE and CIJE covering the period February 1974 to March 1976 were stored on direct access storage. To do this, and also to store the CICS programs, the STAIRS programs, user queries, stop words, sign-on codes and passwords, a single disc pack of 200 million characters capacity was required (IBM 3330, Model 11). It was estimated that up to three years of the data base could be stored on this medium before a second disc would be required. The Project used a share of IBM's 370/158 computer in Canberra, and this share was determined by the computing system in accordance with the demands placed upon it by Project and by other users located all over Australia. The message handling system, CICS, managed the interface between STAIRS and the operating system (OS/VS1), and controlled communication between STAIRS and the terminals (IBM 3275 display units and IBM 3284 printers).

Since the main objective of the Project was to evaluate the use of STAIRS, rather than the computing efficiency of STAIRS, this report does not describe STAIRS in any detail. Interested readers are referred to the appropriate IBM documentation for an exhaustive description:

> IBM. Storage and information retrieval system/ virtual storage (STAIRS/VS): Program reference manual (SH12-5400-1). Second edition, August 1974 - plus updates. 237p.

The following brief outline is taken from the introduction to the program reference manual:

"The IBM System/370 Storage and Information Retrieval System/Virtual Storage (STAIRS/VS) is a multi-user system for the storage and retrieval of documents. It offers the user a variety of resources for the creation and maintenance of data bases and, especially, for the retrieval and output of documents, to be performed either in an on-line dialog using terminals, or in batch mode. Documents, which may contain textual and/or formatted data, are retrieved by means of queries to be formulated by the user in extended Boolean logic and/or comparison operators".

Four separate data bases were available to users:

(i) latest month of RIE

(ii) latest month of CIJE

(iii) two years of RIE

(iv) two years of CIJE

Data bases (iii) and (iv) included data bases (i) and (ii) respectively. The annual cost of updating the data base was estimated at $6000.

Users specified at the beginning of each terminal session which data bases they wished to search. They then had the option of developing effective queries on one data base (it was most efficient to do this on a small data base), saving their queries, then changing to another data base for searching with the newly-developed queries.

4. The surveys within the Project

Two of the primary objectives of the Project were directly related to the operations of the Education Information Service. One called for an investigation of:

> "the extent to which an on-line system such as STAIRS is able to meet the data processing requirements of the Education Information Service for the following five years";

and the other for an analysis of:

> "the costs and benefits of on-line education information retrieval compared with off-line education information retrieval"

In the light of these objectives, it was decided to conduct two small surveys - one to compare the methods of data processing currently available to the Education Information Service (EIS), and the other to investigate the feasibility of using STAIRS for off-line information retrieval.

A. STAIRS as the basis for EIS data processing

During the one month's extension of time granted to the On-line ERIC Project, the feasibility of using STAIRS for the EIS's batch processing operations was investigated in conjunction with IBM. The EIS currently uses another IBM software package, TEXTPAC, for its off-line processing.

Sample

10 profiles were selected from those currently being processed by the EIS. The profiles encompassed all facets of one particular field of interest (school/community relationships) and included author, descriptor and free text searching.

Procedure

The profiles were modified slightly to conform with the STAIRS input format and then grouped in two ways: the first grouping consisted of profiles more or less in their original TEXTPAC formats - each profile statement was allocated a separate STAIRS query number; the second grouping combined queries as much as possible, and profiles that originally contained more than ten statements were reduced to less than five. The purpose of grouping the profiles in two different ways was to investigate the difference made to processing costs.

Both groups of queries were input on-line against only a month of data (March 1976 CIJE) - this was to ensure that data entry costs would be kept to a minimum. The queries were saved and then a week later were executed in batch mode against the latest month's data from RIE and CIJE. Both groups of profiles were searched against RIE and CIJE. The input and batch mode execution operations both involved system start-up and shut-down.

Costs

	INPUT COST	COST/ PROFILE	EXECUTION *COST (CIJE)	COST/ PROFILE	EXECUTION *COST (RIE)	COST/ PROFILE
GROUP 1 (fragmented)	$77.12	$4.70	$80.75	$5	$86.30	$5.60
GROUP 2 (combined)	$50.85	$2.00	$75.26	$4.50	$100.32	$7

* includes $20-$30 start-up/shut-down

The costs outlined in the above table were based solely on data processing. Hardware rental was not included because these costs would already be covered if the EIS offered a demand search service. Similarly, costs for both ERIC tape conversions and the addition of new tapes to the data base were deliberately excluded.

Conclusions

Even though the entry costs of between $2 and $4.70 per profile were high, they would only be incurred each time a profile was changed. Although entry costs were calculated against a real data base, it was possible to establish a very small dummy data base containing a minimal number of records. There would be a significant difference in the cost of profile entry against such a data base because of the relatively small number of records to be searched.

The cost of processing per profile per month (both RIE and CIJE are searched each month) using STAIRS varied between $10.60 and $11.50, depending on the precision of profile construction. As the cost for TEXTPAC is only $5.50 per profile per month, a change to STAIRS would involve up to an extra $1500 per month for 250 profiles.

Even the obvious advantages of STAIRS, such as greater speed and reliability, cannot outweigh the additional costs that would also have to be met if CICS/STAIRS was to be run on a non-IBM service bureau computer.

Though off-line STAIRS is far too expensive for the EIS to consider implementing at present, a re-examination in 12 months time will be worth-while because of the possibility of reduced costs for data entry and shared use of STAIRS by cooperating institutions.

B. Comparative information retrieval costs

Search methods

Three methods are currently available for interrogation of the ERIC data base: manual, batch and on-line. In comparing these methods, two main considerations were taken into account:

(i) cost (of staff time, data processing and presentation of results), and

(ii) relevance of retrieved citations.

Sample

Three people familiar with the scope of the ERIC data base each provided three sample queries. From these nine, three were selected: one straightforward; one involving several overlapping areas of interest; and one which was not adequately covered by the Thesaurus of ERIC descriptors. The selected queries were:

Query 1 (simple)

Examples of high school curricula in home economics

Query 2 (involving two or more terms from the Thesaurus of ERIC descriptors)

Segregation versus integration of mentally handicapped children

Query 3 (outside the scope of the Thesaurus of ERIC descriptors)

Camp schools

The sample was necessarily small because the on-line searching component of the study had to be completed before the end of the On-line ERIC Project. Most of the time available was already allocated to members of the pilot user group.

Procedure

The March 1976 issue of RIE was chosen as the basis for the searches mainly because it was the next tape to be processed for the SDI service.

The cost of search tools - the Thesaurus of ERIC descriptors and the data base itself were not taken into account because these were common to all three search methods.

Method

(i) Manual

The Thesaurus of ERIC descriptors was consulted for each of the three queries to determine relevant terms and these were searched in the subject index in the back of RIE. Because the index is limited to major descriptors, relevant citations were possibly missed. Very few citations were found: one for Query 1, five for Query 2, and one for Query 3.

Each phase in the manual search procedure was timed and later costed. Query 1 took 12 minutes to complete, Query 2, 16 minutes, and Query 3, eight minutes. These were costed at Librarian Class 2 salary rates. The results are given in the following table:

MANUAL PROCESSING

QUERY	TIME TAKEN TO CONSULT THESAURUS AND MARCH RIE (in minutes)	COST OF STAFF TIME (at L2 salary)	NO. OF CITATIONS RETRIEVED	COST OF PHOTOCOPYING
Query 1				
Examples of high school curricula in home economics	12	$1.32	1	5c.
Query 2				
Segregation vs integration of mentally handicapped children	16	$1.76	5	25c.
Query 3				
Camp schools	8	88c	1	5c.

(ii) Batch

Using TEXTPAC, each query was translated into a profile containing terms from the Thesaurus of ERIC descriptors. Much more flexibility was possible compared with the manual method because the searches did not have to be limited to the descriptor fields of the citations. The profiles were typed, input as part of a normal EIS profile update run and then searched against the March 1976 RIE tape. Both of these operations were performed on the computer at IBM. Results were printed as part of the normal EIS print run on the National Library's IBM System 3 computer.

The time taken to complete each part of the batch search operation was noted and later costed. The final costing included totals for staff time (at Librarian Class 2 and Clerical Assistant Grade 3 levels), and data processing time. The complete breakdown is as follows:

BATCH PROCESSING

QUERY	PROFILE CONSTRUCTION (in minutes)	COST (at L2 salary)	TYPING AND INPUT (in minutes)	COST (at CA3 salary)	COST/PROFILE FOR SEARCH RUN	NO. OF CITATIONS RETRIEVED	COST OF PRINTING
Query 1							
Examples of high school curricula in home economics	7	77c	3.75	23c	$3	4	Not Available
Query 2							
Segregation vs integration of mentally handicapped children	10	$1.10	5	30c	$3	2	N/A
Query 3							
Camp schools	5	55c	2.5	15c	$3	10	N/A

(iii) On-line

This search method again involved the use of the IBM computer but the data base was interrogated using STAIRS.

In terms of search strategy, STAIRS permitted the same degree of flexibility as TEXTPAC because searching did not have to be limited to any particular fields of the citations. The Thesaurus of ERIC descriptors was used to formulate the search strategy before the terminal session commenced. However, in viewing output, STAIRS permitted much greater flexibility than TEXTPAC because retrieved citations could be browsed on-line and the search strategy amended accordingly. The results were printed off-line at IBM in Canberra.

Both search preparation time and the time taken to execute the search on-line were noted and costed. Data processing and printing costs have also been included:

ON-LINE PROCESSING

QUERY	PREPARATION TIME (in minutes)	COST (at L2 salary)	TIME ON-LINE (in minutes)	COST ON-LINE (at L2 salary)	COST OF DATA PROCESSING	NO. OF CITATIONS RETRIEVED	COST OF PRINTING
Query 1							
Examples of high school curricula in home economics	4	44c	15	$1.64	$18.60	3	27c
Query 2							
Segregation vs integration of of mentally handicapped children	5	55c	10	$1.10	$12.40	10	90c
Query 3							
Camp schools	2	22c	7	77c	$8.68	9	81c

Search results

The people who had supplied the queries were sent photocopies of all retrieved citations. Although computer print-out is the normal output from both batch and on-line search systems, photocopies were sent to preserve the anonymity of the search procedure used in each case.

Citations retrieved and relevance judgements are incorporated in the table below:

COST and RELEVANCE

	MANUAL			BATCH			ON-LINE		
	Query			Query			Query		
	1	2	3	1	2	3	1	2	3
Total cost	$1.37	$2.01	93c	$4	$4.40	$3.70	$20.96	$14.95	$10.70
No. of citations retrieved	1	5	1	4	2	10	3	10	9
Relevance	100%	60%	100%	100%	50%	50%	100%	10%	66%
Cost/citation	$1.37	40c	93c	$1	$2.20	37c	69c	$1.50	$1.17
Cost/relevant citation	$1.37	67c	93c	$1	$4.40	74c	69c	$14.95	$1.78
Average cost/ citation		90c			$1.52			$1.12	
Average cost/ relevant citation		99c			$2.05			$5.81	

Results

(i) Cost

The table above shows the total cost involved in processing the three queries using each of the search methods. The figures took staff, data processing and printing costs into consideration.

There was very little difference between the average cost per retrieved citation for each of the processing methods: a range of 90c to $1.52. However there was a marked difference in the cost per method for the individual queries: Query 1 varied between 69c and $1.37, Query 2 between 40c and $2.20, and Query 3 between 37c and $1.17. But no pattern emerged from the variation: the cost did not change consistently with the method of processing.

(ii) Relevance

Relevance varied depending on the method of processing for Queries 2 and 3. In each case the manually derived results were judged more relevant. Even though the relevance ratio was higher for these results, more citations were retrieved using TEXTPAC and STAIRS. The citations retrieved for Query 1 were 100% relevant in each case.

Again no pattern emerged for the cost per relevant citation according to the method by which each query was searched. However a quite dramatic difference was apparent when the average cost per relevant citation for each method was compared. The result varied from 99c for the manual searches to $5.81 for the on-line searches. This figure is high because only 10% of citations retrieved for Query 2 using the on-line method were judged relevant (even though the person concerned stated that all were relevant to the subject field).

Conclusions

Unfortunately the results of this study must be deemed inconclusive because of the high degree of variance exhibited. This variance can be attributed to the extremely small sample chosen for reasons already outlined. Although no real conclusion could be drawn regarding either the costs of the three search methods or the relevance of the results obtained, it was quite obvious that TEXTPAC and STAIRS offered far more flexible methods of searching than the manual approach.

To make comparison worthwhile, any future study must be based on a much larger sample.

Comparisons with the Project

It is interesting to compare the results of this study with figures derived from typical On-line ERIC terminal sessions.

During three typical sessions, costs of 97c, $1.09 and $1.36 per citation retrieved were calculated. These figures are comparable to the average cost per citation of $1.12 determined during this study.

IV. USER RESPONSE

The user group represented the largest series of unknown variables in the Project, and users' reactions to the on-line facility are the basis of this report. Apart from a log of system failures and the computer listings produced for each session, the main source of data was completed terminal users' questionnaires. Before each session "serious users" were asked to describe briefly the topics they intended searching, to indicate whether they intended to use STAIRS for specific or exhaustive searching, and to estimate the number of citations they expected to retrieve. At the end of their sessions at the terminals they were asked to answer more questions relating to their use of the system and to the value of their results. The analysis in this section is based upon 261 completed or partially completed questionnaires from 34 users in Canberra (81 responses), 34 users in Melbourne (106 responses), and 46 users in Sydney (74 responses).

In case it is thought that the statistical analysis of user responses confers some sort of objective status on the results themselves, it must be emphasised that the questionnaire answers, except for the answers to factual questions (how much? how long?), are essentially subjective. Value judgements are inevitable where value is being assessed, and while the preceding section of this report dealt with the efficiency of the system, this section concerns itself more with users' judgements of the effectiveness of on-line ERIC.

Since the quantity of data collected was quite small compared with the quantity of data collected for the ERIC Research Project, simple statistical techniques have been used. Tables show numbers, averages (arithmetic mean), percentages, and where applicable, modes (most frequent value), medians (middle value) and ranges. These last three are presented to give a more accurate description of results than is possible by showing only the arithmetic mean. In all cases, relevance judgements were made by users in accordance with their own, unknown, standards.

1. User group origin

Two thirds of the user group were employed by universities or by colleges of advanced education (henceforth referred to as "colleges").

NO. OF USERS	CANBERRA	MELBOURNE	SYDNEY	TOTAL AS A % OF TOTAL USERS
from federal departments and agencies	18	1	0	17
from state departments and agencies	0	4	4	7
from industrial and commercial organisations	1	3	1	4
from associations and voluntary organisations	2	3	0	4
from state library authorities	0	1	0	1
from universities	2	9	36	41
from colleges	11	13	5	26
TOTALS	34	34	46	100%

(Categories are taken from Development of resource sharing networks, Canberra, NLA, 1975)

It was the aim of the project team that the user group should represent as closely as possible the actual consumers of education information, and given the slant of the ERIC data base towards educational research rather than immediate classroom application, it appears that for a population of only 114 users this objective was achieved. With the exception of teachers, the user group appears to reflect a representative proportion of those engaged in educational research.

The preponderance of federal government employees in the Canberra group can be interpreted in terms of the administrative role of the national capital where the government is easily the largest single employer. Sydney shows a much higher proportion of university users than Melbourne, but Melbourne shows the highest proportion of college users. Considering the major role of colleges in the Victorian education system, this is not surprising. Nor is the number of university users in Sydney remarkable when the strength of Macquarie University's Education Faculty is taken into account.

2. Subject coverage

Question 1 asked users: "Please briefly describe the topic you intend searching". The intention of this question was to determine the most active areas of educational research and ultimately to use the collected data as a guide for the development of specialised education information services. The answers appear in the table below:

Subject coverage (number of topics covered)

	TOPIC	CANBERRA	MELBOURNE	SYDNEY	APPROX % of 261 TOPICS SEARCHED
1	ABILITIES	1	3	3	3
2	ADMINISTRATION	2	1	2	2
3	ARTS	1	1	0	1
4	ATTITUDES	1	6	2	3
5	BEHAVIOR	4	3	0	3
6	BIOLOGY	1	2	2	2
7	COMMUNICATION	0	1	1	1
8	COUNSELING	0	1	0	0.5
9	CULTURE	0	2	1	1
10	CURRICULUM	1	5	2	3
11	DEMOGRAPHY	0	4	0	1.5
12	DEVELOPMENT	1	1	4	2
13	EDUCATION	14	15	23	20
14	EMPLOYMENT	0	1	0	0.5
15	ENVIRONMENT	2	0	0	1
16	EVALUATION	6	2	1	3
17	EVAL. TECHNIQUES	0	1	1	1
18	EXPERIENCE	3	0	1	1.5
19	FACILITIES	11	6	1	7
20	FINANCE	3	0	0	1
21	HANDICAPPED	4	0	3	3
22	HUMANITIES	1	0	0	0.5
23	INSTRUCTION	3	8	3	5
24	INSTRUCTNL. PROG. DIVS	1	0	0	0.5

TOPIC	CANBERRA	MELBOURNE	SYDNEY	APPROX % of 261 TOPICS SEARCHED
25 LANGUAGE & SPEECH	0	5	1	2
26 LEARNING & COGNITION	0	7	1	3
27 LIBRARY MATERIALS	1	0	0	0.5
28 LIBRARY SCIENCE	3	6	4	5
29 MATHEMATICS	0	0	1	0.5
30 OCCUPATIONS	0	1	0	0.5
31 ORGANISATIONS	0	0	1	0.5
32 PERSONNEL & GROUPS	4	2	3	3
33 PSYCHOLOGY	1	5	3	3
34 PROGRAMS	0	0	1	0.5
35 READING	2	3	1	2
36 RESEARCH	2	2	3	3
37 SCHOOLS	0	0	1	0.5
38 SOCIAL SCIENCES	4	0	1	2
39 SOCIOLOGY	1	2	0	1
40 STANDARDS	0	1	0	0.5
41 TECHNIQUES	3	0	0	1
42 TESTS	0	9	3	4
TOTAL	81	106	74	99.5%

(Subject categories are taken from the Thesaurus of ERIC Descriptors, 5th edition, Macmillan Information, New York, 1974. pp.313-6).

The scatter of subjects is quite obvious, as is the fact that users did not restrict their subject interests to those topics directly concerned with educational research. The four topics "education", "facilities", "instruction", and "library science" comprised 37% of the total number of topics, with the remaining 63% embracing another 38 topics. Easily the largest category was "education" (20%), and this division included general education concepts as well as specific types of education such as cooperative education, educational improvement, art education, mathematical education, vocational education, etc.

In the opinion of the project team the user group, although numerically small, represented a surprisingly large number of educational subjects. The concentration in the category "education" was probably a reflection of the generalised description given that category by the producers of the Thesaurus.

3. User expectations

The second and third pre-session questions were designed to discover what users expected of the information retrieval system. Answers to these questions could then be compared with answers to post-session questions to determine the extent to which the data base provided predictable results. Previous experience with the ERIC system would obviously enhance the accuracy of the estimate.

Question 2: "Do you intend to use STAIRS:

- to find the latest references to your topic?

- to conduct a fairly exhaustive search of the data base on your topic?

- for some other reason? Please specify:"

TYPE OF USE	CANBERRA	MELBOURNE	SYDNEY
% latest reference	36	38	42
% exhaustive search	80	83	84

The answers showed a clear preference for using the system in its most helpful mode: as an effective means of retrieving a relatively small number of citations from the whole of a relatively large file. Users showed little inclination to search for latest citations only. It must be remembered that this was for most users the first occasion on which they were able to conduct a thorough search of two years of ERIC data, and naturally they would tend to conduct retrospective searches - which would normally be time-consuming, rather than current awareness searches for latest citations only - which they could do quickly anyway by searching RIE and CIJE.

Only 18 "Other reasons" were given, and these ranged from "To support ongoing research project" through "To sample the coverage of Australian material in ERIC" to "To understand the resource information system itself" and "To refine existing SDI profile".

Question 3 was a follow-up to Question 2: "The data base contains almost 70,000 citations. How many do you expect to retrieve in this terminal session?"

EXPECTED RETRIEVED CITATIONS	CANBERRA	MELBOURNE	SYDNEY
Average	64	53	156
Median	30	30	50
Mode	50	20	100
Range	3-1000	7-325	5-3000

Since a few users expected large numbers of citations, and their expectations tended to distort the average figures, medians, modes and ranges have been included. The median figures are probably the most reliable indications of the overall situation. The inflated average value for Sydney was caused by one user who expected more than 3000 citations in answer to a query on preschool education, and two other searches for two different users on teacher education and career education were each expected to uncover more than 1000 citations. In Canberra and in Melbourne users were a little more cautious, and as it turned out, their results were quite close to their estimates. The three

large Sydney searches retrieved 140, 600 and 19 citations respectively, of which 35, 300 and 19 items were judged relevant.

4. Interruptions

The first post-session question was concerned with the incidence of system failure. Question 4 read: "Was your terminal session interrupted by any failures in the computing system or in the terminal equipment?" Users were asked to reply "yes" or "no".

	CANBERRA	MELBOURNE	SYDNEY
% of YES answers	31	36	41

Answers here confirmed the fact that there were more problems with the Sydney installation than with the Canberra or Melbourne centres, but it was difficult to determine the nuisance value of the interruptions. While users were annoyed if their sessions were interrupted, or even if the response time slowed appreciably - particularly if they were in the final steps of refining their search strategies - on the whole they accepted system limitations calmly. Maybe they were suffiently familiar with data processing systems to realise that no implementation is perfect, or perhaps they accepted system failures easily because they knew they were participating in an experiment. In any case, a similar incidence of failure in an operational system would be too high to be acceptable to users or system developers.

5. Users' assessments of their results

The next question was a straightforward request for a subjective evaluation. Question 5: "Were the results of your on-line search:

- of major value?
- of considerable value?
- of minor value?
- of no value?"

% VALUE	CANBERRA	MELBOURNE	SYDNEY
major	29	20	30
considerable	58	61	60
minor	13	18	10
no	0	1	0
TOTAL	100%	100%	100%

Answers here are really encouraging for the future development of on-line systems. With over 80% of completed questionnaires indicating that users consider such services of major or considerable value, it was clear that user acceptance of the products of technology applied to information retrieval was quite high. When it is remembered that the On-line ERIC Project delivered the first on-line, user-operated information service in this country, these ratings are extremely favourable. They certainly emphasise the importance of the development of on-line services.

It remains to say by way of qualification that information consumers in this country are by comparison with their North American and European counterparts, relatively inexperienced in their evaluation of information systems. This is inevitable since they have not had the opportunity to use and compare sophisticated systems. Therefore they can probably be expected to exhibit a tendency to rate highly any system which improves information delivery in their fields - if only on the basis that something is better than nothing. Notwithstanding this qualification, the results are encouraging for the developers as well as the users of on-line information retrieval systems.

6. On-line performance

Question 6 asked: "How long did your search take?"

TIME IN MINUTES	CANBERRA	MELBOURNE	SYDNEY
Average	40	33	67
Mode	30	30	90
Median	35	30	65
Range	10-80	5-100	12-180

Over the three months of the Project there were 81 scheduled sessions for Canberra users, 106 for Melbourne users and 74 for Sydney users. The table above mirrors this level of activity: the longer the session time, the smaller the total number of sessions. When designing the study the project team expected that the time taken for users to complete their searches would decrease as they became increasingly familiar both with their information needs and with STAIRS. While this pattern was confirmed in Canberra and in Melbourne where 60 minute sessions became 30 or 20 minute sessions for subsequent searches, this reduction in searching time was more closely attributable to tight scheduling by project co-ordinators than to increases in user efficiency. It seems that the scheduled three or more sessions for "serious users" gave insufficient time for developing complete familiarity with the system, and users in all three locations often complained of insufficient terminal time. What happened was that users' information

requirements expanded to fill the available terminal time once they had satisfied themselves that the system worked. The Project offered to a small group of users for a limited period a large file of accessible items. It was not surprising that users sought to make the most of the opportunity, and searched as many topics as possible within the allocated time. For the experienced staff at the National Library, searches rarely took longer than 20 minutes, with 10 minutes as an average for uncomplicated, single-topic queries.

On-line information retrieval is fast. The speed with which the computing system can respond to instructions (usually 3-4 seconds) is simply astonishing to the uninitiated. More users requiring more information seems a reasonable prediction for the field of information delivery, and in the development of on-line systems resides the potential to meet high-volume demand.

There was no doubt that increased familiarity with the computing system brought more efficient use of the system, and to this end, throughout the Project, co-ordinators either actively assisted users by sitting with them at the terminals, or they sat within earshot, ready to deal with questions before they became problems. Without exception, users found this the most satisfactory (and secure) operating environment.

Question 7 was directed towards users' search results: "How many citations were retrieved by your final search strategy?"

NUMBER OF CITATIONS	CANBERRA	MELBOURNE	SYDNEY
Average	70	81	116
Median	39	35	60
Range	1-1300	2-1000	0-950

Given the interactive searching capability of STAIRS, these values seem high. It was possible for users to interactively refine their queries until they had found only the relevant and all the relevant citations, and if this had happened, the figures would have been lower. It was certainly the case that three terminal sessions were not sufficient for users to become completely familiar with the system. However, the answers to Question 2 indicated that users were more interested in retrospective than in current awareness searches, and this attitude probably accounts for the high figures in the table above.

When the number of retrieved citations is compared with the number of expected retrieved citations, a close correlation is evident.

AVERAGE NUMBER OF CITATIONS	CANBERRA	MELBOURNE	SYDNEY
Expected retrieved	64	53	156
Actually retrieved	70	81	116

In Canberra the number of retrieved citations was 109% of the number of expected retrieved citations; in Melbourne it was 153% - the largest variation; and in Sydney it was 74%. This finding tends to support the contention that users were not completely unfamiliar with the contents of the data base nor with the volume of literature in their specialist fields. It is interesting to note that only in Sydney, where university and college staff comprised 89% of total users, was the number of expected retrieved citations higher than the number of retrieved citations. In Canberra (38% of users from universities and colleges) and in Melbourne (65% of users from universities and colleges) the number of expected retrieved citations was lower than the number of retrieved citations.

Given answers to Questions 6 and 7 it was possible to calculate the time taken to retrieve one citation by dividing the number of citations retrieved into the time in minutes taken to retrieve those citations. The results underscore the speed of on-line information retrieval:

TIME IN MINUTES TO RETRIEVE ONE CITATION	CANBERRA	MELBOURNE	SYDNEY
Average	2.9	0.45	1.5
Median	1	0.87	0.9
Range	0.02-40	0.07-15	0.1-15

The interesting figure here is the median. At around one minute to retrieve one citation, the three locations operated at roughly equal retrieval speeds - and at speeds which could not be consistently matched by either off-line or manual systems.

Naturally, the number of retrieved citations is by itself a poor indicator of the efficiency of the system and must be taken in conjunction with the number of relevant retrieved citations. Whereas Question 7 asked how many citations were retrieved, Question 8 asked:

"How many of these were relevant to your topic?"

Average, median and range figures here may be misleading because the number of relevant retrieved citations is really only significant when compared with the number of retrieved citations. Precision figures (the proportion of relevant to retrieved citations) have been calculated for each individual session on the basis of users' assessments of the relevance of the items retrieved.

PRECISION RATIO (%)	CANBERRA	MELBOURNE	SYDNEY
Average	70	65	68
Mode	100	100	100
Median	76	79	73
Range	7-100	0-100	0-100

Despite the fact that these are high precision figures, they can only be taken as a rough guide to the performance of the system. To turn them into reliable figures they need to be compared with recall ratios (the proportion of relevant citations retrieved to the total number of relevant citations). But to have determined the total number of relevant citations for a particular query would have involved the project team in sampling studies far more complex and time-consuming than was warranted by the advantage of knowing the answer.

The available experimental evidence suggests that recall and precision tend to vary inversely; that the higher the number of citations retrieved the lower the likelihood that individual citations will be relevant. Given this known relationship, and the fact that recall and precision are purely subjective measures, no recall figures were determined. In any case, how does one evaluate a search which retrieves zero citations? A

user may consider this result highly significant if he/she is beginning a new project and wants to survey the existing literature. In this case a nil result may be very encouraging indeed!

The earlier ERIC Research Project, which was similar in conception to the On-line ERIC Project (except that it was based upon an off-line system, TEXTPAC), resulted in precision ratios of 59% for current awareness searches and 45% for retrospective searches. The On-line Project results were significantly higher.

But there is more to it than this. Notwithstanding the fact that the precision ratios were quite good, an on-line system such as STAIRS has a theoretical capability of 100% precision. Consider a typical terminal session: let us assume that the user is interested in psychological studies of the extent to which television viewing leads to anti-social behavior, in particular aggression or violence, in children. The user keys in the terms "ANTI SOCIAL BEHAVIOUR" (1). The system responds with a tally of the number of occurrences of the phrase in the data base, and also with a count of the number of citations in which the phrase occurrs. The user then enters the next query "AGGRESSION OR VIOLENCE" (2) and the term and document counts are again displayed on the screen. Then the user enters "CHILDREN" (3). The next step is to combine the query in the form (1 OR 2) AND 3 - this becomes query 4. When query 4 is combined with "PSYCHOLOGICAL STUDIES" in query 5, the result is displayed on the screen. The user may then ask STAIRS to display the answers to his/her final search statement,

and the document citations themselves appear. He/She can then select the relevant from the displayed citations - 100% precision. Such powerful searching capability is simply not available with anything other than an on-line system.

With users' relevance judgements for retrieved citations it was possible to determine how long it took to retrieve a single relevant item, and to compare that value with the time taken to retrieve an item:

LOCATION	AVERAGE TIME IN MINUTES TO RETRIEVE	
	ONE CITATION	ONE RELEVANT CITATION
CANBERRA Average	2.9	3.9
Median	1	2.14
Range	0.02 -40	0.11-40
MELBOURNE Average	0.45	3.8
Median	0.87	2.1
Range	0.07-15	0.07-30
SYDNEY Average	1.5	3.2
Median	0.9	1.5
Range	0.1-15	0.11-75

The median figures for the time taken to retrieve one relevant citation (2.14, 2.1 and 1.5) are surprisingly close, reflecting a uniformly high level of performance at the three centres. It is interesting to note that the medians for relevant citations are roughly double the medians for retrieved citations - and this is probably a reflection of the way users used the system. That is, they were not particularly concerned to retrieve only the relevant citations but were satisfied if their results included at least the relevant citations.

Question 9: "How many of the relevant citations were completely new to you?"

AVERAGE	CANBERRA	MELBOURNE	SYDNEY
number of relevant retrieved citations	34	36	81
number of new citations	32	33	77
% of new to relevant citations	94	92	95

These results, especially the percentage figures, point to the facts that the ERIC data base is a relatively little-used source of educational information, and moreover, that it is a useful source - more than 90% of the

relevant items retrieved were unknown to users. Answers to this question can be interpreted as a vindication of the National Library's decision to make ERIC information services available in Australia.

As a final comparison in this section:

AVERAGE	CANBERRA	MELBOURNE	SYDNEY
time in minutes to retrieve one citation	2.9	0.45	1.5
time in minutes to retrieve one relevant citation	3.9	3.8	3.2
time in minutes to retrieve one relevant new citation	4.1	3.9	3.8

Here the range has narrowed still further (only 0.3 of a minute difference, on these average figures), indicating that throughout the project the Canberra, Melbourne and Sydney terminals operated at much the same level of efficiency - despite the difference in numbers of retrieved citations.

Questions 10 and 11 asked users about printing their results: Question 10: "How many citations did you print at the terminal?" and Question 11: "How many citations did you have printed in Canberra?"

AVERAGE	CANBERRA	MELBOURNE	SYDNEY
number of citations retrieved	70	81	116
number of relevant citations retrieved	34	36	81
number of citations printed at the terminal	0.2	0.4	0.7
number of citations printed off-line	60	60	83

The interesting conclusion here is that users tended to print at least all the relevant citations and quite often they printed all the retrieved citations - presumably to recheck their relevance judgements. In 51 out of 261 sessions, users printed more than 100 items, and one user printed a 1300-item bibliography on pupil assessment and pupil-teacher relationships. Given the fact that it was very easy for users to print large result files, the quantity of printing done seems to bear out the point made above that users really took up with alacrity the opportunity to use the on-line system for tasks they had never previously been able to perform.

And it appeared that users had not previously been able to produce bibliographies tailored to their specific interests. Very little printing was carried out on the terminals attached to the printers because when the printer was operational the screen was inactive. In general users preferred to issue a simple command at the terminal and have their results printed off-line, usually overnight in Canberra, and then receive them by express delivery.

7. ON-LINE EFFECTIVENESS

Questions 12-17 were concerned with users' assessments of the effectiveness of the system and their comments on the usefulness of the training materials.

Question 12 was really another version of Question 5, except that it called for a simpler answer. Given that most respondents to question 5 indicated that their search results were of major or considerable value, the answers to question 12 were predictable. "Were your search results worth the time and effort spent finding them?

YES/NO/HARD TO TELL"

	CANBERRA	MELBOURNE	SYDNEY
% Yes	86	77	93
% No	1	1	0
% Hard to tell	13	22	7
TOTAL	100%	100%	100%

The two "NO" answers came from particularly unfortunate users - one who retrieved 45 citations, none of which were relevant, and one who retrieved two citations in 30 minutes, only one of which was relevant. The "HARD TO TELL" answers usually reflected low levels of relevance for retrieved citations. But in general, the answers to this question corroborated the answers to question 5, and confirmed the fact that users found the system useful and relatively easy to use.

Question 13 related to user documentation: "Did you find the users' manual Easy steps to STAIRS:

- very useful?
- useful?
- no use at all?"

	CANBERRA	MELBOURNE	SYDNEY
% very useful	48	39	57
% useful	52	60	36
% no use at all?		1	7
TOTAL	100%	100%	100%

Some Sydney users found the users' manual more useful after than before terminal sessions, but since the manual was intended as a guide to the system and as a reference for command formats and for instructing STAIRS to produce the desired outputs, this negative response is puzzling. All users were sent a copy of the manual several weeks before their scheduled sessions and it is to be expected that some would read it more carefully than others. Overall responses justified the effort the project team put into producing the manual: in fact the team went to considerable lengths to ensure that the manual was easy to use, comprehensive and interesting. It is included as the appendix to this report.

The other major training aid produced was a tape/slide sequence illustrating a typical search. Originally envisaged as a video tape recording, it proved impossible to eliminate the stroboscopic effect produced in a video tape of a video terminal, and the tape/slide format was substituted. The production of the tape/slide sequence was very time-consuming and involved co-ordinating the computing system, photographers, script readers, tape recorders, and then producing two complete duplicates. The project team, having no previous experience in this field, felt that the end result was less than perfect and the imperfections were noticed by some users whose most consistent complaint was that material was presented too quickly.

"Did you find the tape/slide sequence <u>STAIR cases</u>:

- very useful?
- useful?
- no use at all?
- didn't hear/see it?"

	CANBERRA	MELBOURNE	SYDNEY
% very useful	36	3	23
% useful	42	59	36
% no use at all	0	33	14
% didn't hear/ see it	22	5	27
TOTAL	100%	100%	100%

The comparatively higher Canberra ratings were probably due to the fact that Canberra had the original version and Melbourne and Sydney the duplicates. Since users were only required to view Easy steps before their first terminal session, for subsequent sessions they checked the "didn't hear/see it" category. There is no doubt that this training aid was well worth producing, and it provided a valuable introduction and a form of familiarisation for those people who had never before used a computing system. Its whole point was to demonstrate that STAIRS was easy to use and the project team felt that a revised version would receive more favourable comment. As it was, the majority of users found it very useful or useful.

Users were then asked: "Did you find the Thesaurus of ERIC descriptors:

- very useful?
- useful?
- no use at all?
- didn't use it?"

	CANBERRA	MELBOURNE	SYDNEY
% very useful	61	62	84
% useful	26	32	16
% no use at all	1	3	0
% didn't use it	12	3	0
TOTAL	100%	100%	100%

Almost 90% of users found the Thesaurus either very useful or useful. Considering that the Thesaurus provided the means by which users could identify their requests to the system in exactly the same manner as the ERIC indexers identified ERIC documents to the system, such a favourable response was hardly surprising.

Question 16 was yet another check on Questions 5 and 12, although from a slightly different approach:

"Please rate the importance of STAIRS/ERIC as an information source:

- extremely useful?
- very useful?
- useful?
- of little use?
- of no use at all?"

	CANBERRA	MELBOURNE	SYDNEY
% extremely useful	65	61	72
% very useful	30	28	20
% useful	4	11	7
% of little use	1	0	1
% of no use at all	0	0	0
TOTAL	100%	100%	100%

This question was a type of summary of all the questionnaire responses, aiming at an overall assessment of the users' experience in using the terminal. As the table indicates, the users thought very highly of on-line ERIC. More than 90% of responses rated the system either extremely useful or very useful and such favourable user evaluation really does speak for itself.

Then in May the project team decided to add another question to the questionnaire asking users to put a dollar value on STAIRS/ERIC:

"While it has not been our policy, nor is it yet our intention to charge for computer-based information services, reduced operating capital leads us to consider at least partial cost recovery. We would appreciate your estimate of a fair price for a service such as on-line ERIC.

In making your estimate please consider the value of rapid access to current developments in your field and the value of the time saved through using an on-line system. The data processing cost of a typical 30 minute terminal session is about $30.

Please indicate below the amount you think a reasonable charge <u>for one terminal session</u>.

- less than $10
- $10 - $25
- $25 - $50
- $50 - $100
- $100 - $200
- more than $200"

	CANBERRA	MELBOURNE	SYDNEY
% less than $10	19	19	0
% $10-$25	44	37	36
% $25-$50	37	40	50
% $50-$100	0	4	9
% $100-$200	0	0	5
% more than $200	0	0	0
TOTAL	100%	100%	100%

These results indicated a certain resistance to paying the total costs of providing the service. This may have been a reflection of personal parsimony, or it may have been evidence of a feeling that employing organisations ought to pay the bill. It seems that users were prepared to pay the data processing costs, but not the total costs which can range up to 500% higher. Perhaps it was a mistake to give some clue on the questionnaire in the form of the cost of a typical session, but without this the estimate would have had to have been made in a vacuum. Sydney, with its high proportion of academics, was obviously readier to pay than Melbourne or Canberra. Many respondents said that such services

should be subsidised in the national interest, but the notion of free information services is pervasive. The question of whether on-line services are worth the cost of providing them seems to have been answered in the affirmative by questions 5,12 and 16 below. When people have been expecting and getting something for nothing, they are understandably reluctant to pay for the same product. On the other hand, respondents may have thought that by indicating a high figure it was more likely that high charges would ultimately be levied.

The one other question on the questionnaire not yet covered was "Do you have any other comments?" In the opinion of the study team these comments warrant reproducing in full in this report. After all, the Project did break new ground.

CANBERRA USERS' COMMENTS

- A first class research tool.
- Efficient use requires remembering STAIRS details.
- Would like more discussion before terminal session, hard to assess value.
- Bit of a lump for non-librarians to swallow.
- Promising.
- Familiarity with the system would enhance its usefulness.
- System is of value and should be continued.
- Potentially much more useful than one session allows.
- Obvious benefits of time saved and efficiency. System enormously attractive.
- Not enough time on the system.

- Not enough time. Slow response. Unable to determine relevance.
- Would like more time to search.
- Plenty of practice is necessary it seems to me.
- Unfamiliarity wastes time.
- Search formulation in printout would be useful. Easy to use.
- Great possibilities for my area of work.
- Extremely useful if developed to relate to Australian publications also.
- Needed continual help. Manual could have been more helpful.
- Would like to be able to use when needed. Heuristic search facilities most important.
- It would be senseless to discontinue the service for economic reasons. Untold help in performing my job.
- A very worth-while exercise and well-organised.
- Three sessions not enough - has been valuable.
- Considerable experience necessary. Best when requests can be narrowly defined.
- I would like to continue to be provided with this service. Extremely valuable.
- In all cases the printed material forwarded to me from my searches has been extremely relevant.
- Too little time to become really at home with the system.
- Widely spaced half hour session not enough.
- Would like to use the system often. Easy way of searching.
- Research in Australia in education would lose a most valuable tool if the STAIRS program was to be discontinued.

- I hope for continuance of on-line access. Essential service of NLA which should be continued and extended.
- "Easy steps" of considerable help having used the system.
- Given the enormous amount of material now available and the shortage of time available for browsing through library shelves, this facility is incredibly valuable to me as a tertiary teacher and educational researcher.
- The system as an information source is of immense value.
- There is no substitute for spending long periods at the terminal.
- A most valuable service.
- Would like to use the system frequently. The searcher changing the query seems to mean higher rate of relevance
- Lack of familiarity with thesaurus meant (that I) haven't fully benefited. The system itself is of immense value.
- The assistance given by the staff at all times was excellent.
- A very good system, should be of great use to educators.
- I received far more citations than expected and had no time to browse them all for relevance.

MELBOURNE USERS' COMMENTS

- Need simulated keyboards for dummy practice.
- Limited to USA.
- Keyboarding takes up a lot of time for complex searches.
- Only 4 relevant citations sighted before "technical breakdown". Technical reliability is essential and should be the norm.
- American orientation.

- Good time saver.
- On-line ERIC, particularly CIJE an extremely useful source of information.
- Very good, but occupying on-line time for printing at the terminal is unfortunate.
- Frequent system failures.
- Need guidance to avoid wasting time.
- Searched only one data base but result satisfactory.
- Would prefer more time to browse and select. Had to use ..print all.
- Greater familiarity and more time needed.
- Doubtful relevance of data base to industrial training.
- I hope this service will continue.
- A message to warn of remaining time left would be helpful.
- Helpful technical assistants are essential for the first session at least.
- A steep rise in the learning curve could be expected after 2-3 hours.
- Need more time to properly evaluate the system.
- Available descriptors brought forth too much irrelevance.
- Not enough time.
- Bewdy!
- More time needed to complete a search.
- Appears to be potentially of considerable value.
- A most valuable system and I would like this service to continue.
- Every family should have one!
- Difficult to assess the value of the information until complete details of the references are seen.
- No generic search capability.
- Insufficient time.

- Need to prepare search strategy essential. Lack of adequate preparation leads to fewer results.
- Excellent when working well. Receiving previous print-out before next session would be an advantage.
- Should be subsidised if necessary. Depends on ease of access and length of session.
- It would be good if the service could be established.
- Slowness of computer response a serious handicap to comprehensive searching.
- 4 different topics searched in an hour.
- Data base more suited to education than librarianship searches.
- This system should be available on a permanent basis.
- Too long a response time.
- Either the user should have regular (weekly) access or should work with expert operators.
- Community affairs marginal subject - difficult to retrieve material.
- It's a pity that this is the last time I can use it in the present run.
- More time to browse required to cut down on print function.
- I would like this service to continue.

SYDNEY USERS' COMMENTS

- Tremendously enthusiastic about this mode of information retrieval.
- An invaluable research aid, which with practice, could recast the nature of literature searching.
- With familiarisation should be of enormous benefit.
- An excellent research advance.

- Took as long as a manual search the first time, but would get quicker.
- Of enormous value.
- Eliminates time wasted in manual cross referencing of a large topic.
- Two breakdowns frustrating, but extremely useful for keeping up to date with overseas developments.
- Some indication, before session, of the need and method for documenting search strategy.
- A staggering, mind-blowing experience. The future, now!
- Prefer ERIC SDI to STAIRS/ERIC.
- "Easy steps" only useful after session.
- Took the place of otherwise very time-consuming literature search.
- Maximum use of the system made possible by very helpful staff.
- Manual did not explain adequately for beginners, how to enter terms, eg. descriptors.
- "Easy steps" and "Stair cases" difficult to comprehend before session. Service of great benefit if available on a permanent basis.
- Became aware of directions and relationships more readily than by other search methods.
- Well worth-while.
- In spite of many combinations of descriptors result was zero. Quite a surprise and in itself a very significant finding.
- A useful start to a comprehensive literature survey on this topic.
- A valuable list of resources from which it has been possible to identify a number of new innovative trends. Most useful.

- Thank you for the opportunity to use the system.
- Citations will be valuable for keeping up to date and will aid further research.
- Didn't expect to find much previous work. STAIRS confirmed this very quickly.
- I found using the system to be totally absorbing and I was particularly happy to be able to use it because I have been engaged in receiving the literature for the Project for the last two months, and using the computer terminal is infinitely superior to working manually with the indexes.
- An invaluable facility for one commencing a major project.
- Excellent service.
- Librarian's assistance invaluable.
- Suggest accompany manual with photographs. Not enough time for complex topic/search. Excellent for single descriptor entry only.
- Very useful system provided there are no mechanical breakdowns.
- Potentially is an enormously powerful research aid.
- I thoroughly believe in the system.
- Marvellous.
- Assistance of Mr. Ray Penn was extremely useful.
- The saving of time and energy is incalculable but experience is necessary - or expert help.
- Material not directly related to this topic would most likely have been related to other aspects of my work.
- Operator extremely competent and helpful.
- Excellent.
- Requires practice.
- I experimented with some Thesaurus terms and this was not at all useful.

V. CONCLUSIONS

1. The earlier investigations of the ERIC Research Project were extended and consolidated by further examination of the applicability of on-line information retrieval techniques to the subject areas of education. The major finding of the On-line ERIC Project was that on-line systems such as STAIRS which facilitate user interaction with the data base and interactive formulation of search strategies, represent a major improvement in information retrieval effectiveness compared with off-line systems.

2. The Project found that an on-line system such as STAIRS has the capacity to meet the retrospective data processing requirements of the Education Information Service for the next five years. Further to this, it was established that direct user interrogation of the data base was a practicable modus operandi for the development of an educational information network. But while STAIRS' information dissemination capabilities were thoroughly investigated during the Project, it was the opinion of the project team that further evaluation of on-line data base creation should be carried out within a network environment. To retain the momentum of the Project, it was felt that this task should be started without delay.

3. The costs and benefits of on-line education information retrieval were compared with the costs and benefits of off-line education information retrieval. While involving higher initial and running costs than off-line sytems, an on-line system such as STAIRS was found to provide significantly improved levels of service to users. The improvement in performance of on-line systems over off-line systems was considered comparable to the improvement in performance of off-line systems over manual systems. Moreover, the establishment of an on-line system with users conducting their own inquiries, with expert assistance, was found to represent the most cost-effective course of action for future network development.

4. On-line access saves users' time and by saving time the system can foster productivity improvements. Most users preferred to use an on-line system themselves to having someone else use an off-line system on their behalf, but until an Australian educational data base is available, changes in information-gathering practices will be difficult to identify.

5. The degree of effectiveness of search formulations written directly by users was found to be almost as high as the effectiveness of search formulations written by trained search analysts. For this reason a decentralised education information network, with either analysts or users assisted by analysts operating the system, was found to be both desirable and feasible.

6. The Project stimulated Australian interest in, and enhanced Australian access to the published results of overseas educational research.

7. It was found that the ERIC data base alone was not sufficient to meet the information needs of Australian educators. Australian material should be as accessible as non-Australian material.

8. On-line systems such as STAIRS are already an effective means, and have the potential to become an efficient means, of improving information delivery in education.

APPENDIX I

EASY STEPS TO STAIRS

ISBN 0 642 99067 0

Education Information Service
National Library of Australia
Canberra 1976

National Library of Australia. Education Information Service.
Easy steps to STAIRS.

ISBN 0 642 99067 0

[1.] STAIRS (Information retrieval system) - Handbooks, manuals, etc. [2.] ERIC - Handbooks, manuals, etc. I. Title.

029.937

EASY STEPS TO STAIRS

LIST OF CONTENTS

	Page
KWIC STEPS TO STAIRS	(vi)
1. INTRODUCTION	1
1.1 STAIRS and the On-line ERIC Project	1
1.2. Before you start	1
1.3 ERIC data	2
1.4 Subject coverage	2
1.5 RIE and CIJE	3
1.6 Tactics	3
1.7 How STAIRS works	3
1.8 Some hints	4
2. TERMINAL EQUIPMENT	5
2.1 Using the terminal	5
2.2 Brightness and contrast	5
2.3 The keyboard	5
2.4 Correcting typing mistakes	6
2.5 Those lights on the screen	7
2.6 What if the system fails to respond?	7
3. OPERATING THE TERMINAL	8
3.1 Signing-on	8
3.2 Changing data bases	9
3.3 Signing-off	9

4. SEARCHING THE DATA BASE 10

4.1 STAIRS modes and functions 10

4.2 What to do when you need help 10

4.3 The SEARCH mode 10

4.4 Search operators 11

4.5 Parentheses 12

4.6 Back-referencing 12

4.7 Knowing where you are 12

4.8 Masking and the ROOT operator 13

4.9 Stopped words 14

4.10 Paragraphs 14

5. WHAT TO DO AFTER THE SEARCH 16

5.1 Looking at your results 16

5.2 Moving around your answer set 16

5.3 Printing your results 16

6. DOCUMENT BACK-UP 18

6.1 Getting hold of ERIC documents 18

7. OTHER FEATURES OF STAIRS 20

7.1 Explanation 20

7.2 Ranking your results 20

7.3 Selecting items 20

8. SAVING QUERIES 23

9. SEARCHING SAVED QUERIES 24

		Page
10.	THE ERIC THESAURUS: a guide to its contents and use	25
	What is it?	25
	Why you should use it	25
	Layout of Thesaurus entries	25
	Notation of the main sequence	26
	Other sections in the Thesaurus	28
	(1) The rotated descriptor display	28
	(2) Descriptor group categories	28
	(3) Descriptor group display	29
	Using the Thesaurus	29

KWIC STEPS TO STAIRS: a Keyword In Context Index to Easy steps to STAIRS.

A. HOW TO		Section(s)
	CHANGE data bases	3.2,8
	COMMUNICATE instructions to STAIRS	2.4
use the	COPY key	5.3
choose	DATA bases	3.1
use the	EXEC function	8,9
use the	HELP function	4.2
	LOOK at your search results	5.1,5.2
	MASK words	4.8
use an	ON-LINE system	1.6,1.7,1.8,4.1
	ORDER ERIC documents	6.1
search	PARAGRAPH classes	4.10
use	PARENTHESES	4.5
	PRINT your search results	5.3
	PURGE unwanted queries	8
use the	RANK function	7.2
use the	ROOT operator	4.8,4.9
	SAVE queries for subsequent searching	8
adjust	SCREEN brightness	2.2
adjust	SCREEN contrast	2.2
turn off the	SCREEN	2.2
turn on the	SCREEN	2.2
use the	SEARCH function	4.3
	SEARCH saved queries	8,9
use the	SELECT function	7.3
	SIGN-OFF at the terminal	3.3
	SIGN-ON at the terminal	3.1
use the	TERMINAL equipment	2.1
use the ERIC	THESAURUS	10
correct	TYPING mistakes	2.4

B. WHAT ELSE		
can you do to improve your	SEARCH formulation	10
can you read about	STAIRS	7.1

C. WHAT IF		Section(s)
you want to get the	FULL TEXTS of retrieved citations	6.1
you need	HELP	4.2
you lose track of your	QUERY numbers	4.7
the system fails to	RESPOND	2.6
you	RETRIEVE too many citations	4.8

D. WHAT IS/WHAT ARE		
the	ALPHABETIC keys	2.3
	BACK-REFERENCING	4.6
the	BROWSE function	5.1
the special	CHARACTER keys	2.3
	CIJE	1.3,1.5
	CURRENT index to journals in education	1.2,1.5
the	CURSOR	2.4
the	DELETE key	2.3
an	ED number	6.1
an	EJ number	6.1
the	ENTER key	2.3
	ERIC	1.3
the	ERIC data base	1.3,1.4
	FORMATTED fields	4.10,7.3
the	FUNCTION keys	2.3
the	INSERT key	2.3
the	KEYBOARD	2.1,2.3
the	LIGHTS on the screen	2.5
	MASKING	4.8
the	NUMERIC keys	2.3
the	ON-LINE Eric Project	1.1
	PARAGRAPHS	4.10
the	PRINTER	2.1
	PUBLICATION type codes	7.3

	Section(s)
the pre-session QUESTIONNAIRE	1.2
the post-session QUESTIONNAIRE	1.2
RIE	1.3,1.5
the RESET key	2.3
RESOURCES in education	1.3,1.5
the ROOT operator	4.8
the SCREEN	2.1
SEARCH operators	4.4
SELECT operators	7.3
the SHIFT keys	2.3
STAIRS	1.1,7.1
STOPPED words	4.9
the ERIC THESAURUS	10
UNFORMATTED text	4.10,5.1

EASY STEPS TO STAIRS

1. INTRODUCTION

1.1 STAIRS AND THE ON-LINE ERIC PROJECT

STAIRS (Storage And Information Retrieval System) is as its name suggests - a system for storing and retrieving information. The On-line ERIC Project, in which STAIRS does the data processing, is designed as an operational evaluation of an interactive information retrieval system. This guide is intended to help you use STAIRS to retrieve references - or more accurately, document citations, from the ERIC data base. (see Sections 1.3-5)

As a self-tutoring system, STAIRS itself is probably your most convenient source of information on STAIRS. You can ask it to conduct a search on your topic and you can ask it how to conduct the search. It enables you to search the ERIC data base, helps you to keep track of how your search is progressing, allows you to refer back to completed steps in your search, and enables you to vary your search strategy. STAIRS is designed to be useful, but to use it productively we think it is important that you read this guide.

1.2 BEFORE YOU START

Before you sit down at the terminal we would like to make a few suggestions and we would like to enlist your support.

- Will you please complete a pre-session questionnaire. This will help us with our evaluation of the system. When you finish at the terminal we would like you to complete a post-session questionnaire.

- Please think carefully about the subject or subjects of your search, and whether you need current or retrospective information.

- Refer to the Thesaurus of ERIC descriptors to select terms for framing your search. The Thesaurus is available from the project co-ordinators. Section 10 describes its use.

1.3 ERIC DATA

The activities of the Educational Resources Information Center (ERIC) are sponsored by the U.S. National Institute of Education. Through a network of 16 clearinghouses, each of which is responsible for a defined subject area, documents are acquired, abstracted, indexed and announced in the monthly journals Resources in education (RIE) and Current index to journals in education (CIJE).

RIE and CIJE are also available in machine-readable form, and the magnetic tapes for RIE and CIJE for 1974 and 1975 make up the data base for this project.

1.4 SUBJECT COVERAGE

There are ERIC Clearinghouses for:

- Career education
- Counseling and personnel services
- Early childhood education
- Educational management
- Handicapped and gifted children
- Higher education
- Information resources
- Junior colleges
- Languages and linguistics
- Reading and communication skills
- Rural education and small schools
- Science, mathematics, and environmental education
- Social studies/social science education
- Teacher education
- Tests, measurement and evaluation
- Urban education

The 1974/5 files are made up of about 70,000 records. Most of these are North American in origin, but some European and Australian items are included.

1.5 RIE and CIJE

RIE is a monthly journal announcing report literature and recently completed research results to the educational community.

CIJE is a monthly comprehensive index to most of the important educational periodical literature. It covers more than 700 journal titles.

1.6 TACTICS

An on-line system responds almost immediately to your instructions. It is therefore possible for you to adopt a trial and error approach to information retrieval. You can develop your search strategy as you go. For example, you could start with a simple query, see how many citations are retrieved, then browse through authors, titles or abstracts looking for terms which could be used to retrieve additional relevant citations or to eliminate irrelevant citations. You could then modify and resubmit your query, repeating the cycle until you feel you have retrieved as much information as you require.

1.7 HOW STAIRS WORKS

The document citations on tape are despatched monthly from Washington. The tapes for 1974/5 have been "inverted" on the computer in Canberra. "Inversion" is a process whereby the locations of words from the citations are recorded in such a way that term occurrences can be directly retrieved - similar to the way in which the index in the back of a book refers to the term occurrences in the text, except that the inverted file is much more comprehensive and records all significant terms, excluding non-significant words such as "and", "at", "the" etc. (See Section 4.9)

The same computer controls the terminals, so if you search for potentially relevant terms, STAIRS can retrieve citations without examining the whole of each record. Since the data base has already been "scanned", and its contents are known to the system, this happens very quickly indeed.

1.8 SOME HINTS

If you have had no previous experience of either on-line systems or on-line information retrieval systems, it does not really matter. Not many people in this country have. We offer the following hints to try to put you at your ease.

- except with malice aforethought, you cannot do any damage to STAIRS. Despite the fact that the writing of such a complex set of programs is quite a feat, the system itself is not fragile.

- By way of a series of messages which appear on the screen, STAIRS will guide your progress through a search of the data base. You can seek further clarification by pressing the HELP key. (Section 4.2)

- Since you will only have to type a few words and press a few keys, you don't have to be an expert typist.

- If you want assistance at any time the terminal is active, you only have to ask the project co-ordinator. Otherwise you will be left to work at your own pace. With the exception of a review of your final search strategy, your work will not be monitored.

- Take your time. Let the system wait for your responses - do not allow it to force the pace.

2. TERMINAL EQUIPMENT

2.1 USING THE TERMINAL

There are three pieces of terminal equipment: a screen; a keyboard and a printer. You use the keyboard to issue instructions and to ask questions. Responses to your instructions and answers to your questions are displayed on the screen. The printer can be used to print your results. (Section 5.3)

2.2 BRIGHTNESS AND CONTRAST

The knob near the lower left hand corner of the screen does three things. If you pull it, you turn on the screen. If you push it, you turn off the screen. Try not to push the knob, because if you do the system will probably have to be reactivated from Canberra, and this could take some time.

Turning the outer part of the knob adjusts the brightness of the display. Turning the inner part of the knob adjusts the contrast between high and normal intensity displays. High intensity is used by STAIRS to highlight query words which cause particular citations to be retrieved. During the sign-on procedure you can adjust the brightness and contrast to suit your preference.

2.3 THE KEYBOARD

The keyboard is made up of:

- 12 special function keys. The names of these keys are displayed on the template which fits around them. They will be referred to below as the HELP key, the SEARCH key, and so on.

- the ENTER key. Depressing this key causes transmission to the computer of instructions which have been typed on the screen by using the keys described below. To communicate your instructions to STAIRS, you can either type them out on the keyboard, then transmit them by pressing the ENTER key, or press one of the 12 special function keys.

- 26 alphabetic keys - exactly as you would find on a typewriter.

- the SHIFT keys - a LOCK key and two SHIFT keys. These work in a similar way to the LOCK and SHIFT keys on a typewriter.

- 10 numeric keys. You can use these to type the digits 0 to 9. By using them with the SHIFT key you can type the special characters shown above the digits.

- Eight other special character keys. Each key has two characters on its face. The lower one is generated if the SHIFT key is released, and the upper one is generated if the SHIFT key is depressed.

- The SPACE bar. This generates blank spaces. By holding the bar down you can automatically generate as many blank spaces as you want.

- The CLEAR key. This clears the screen.

- The INSERT, DELETE and RESET keys and four keys which govern the position of the cursor or movable indicator. (The cursor appears on the screen as a horizontal line about 3mm long. It defines the place on the screen where your next instruction will be entered.) All these keys are used for error correction and their uses are described below.

- Several other keys which are not used with STAIRS.

2.4 CORRECTING TYPING MISTAKES

Instructions typed at the keyboard are not transmitted to the computer until the ENTER key is depressed. If you make a mistake before using the enter key you can use the cursor control key (with the arrow pointing from right to left) to position the cursor under the first incorrect character, and then retype the instruction from that point.

Where you detect an error after you have gone on typing the DELETE and INSERT keys are useful. The DELETE key deletes characters without leaving any gaps. The INSERT key allows you to insert characters without losing or over-typing previously typed characters. You can then use the RESET key and return to your typing. You will need to remember to use the cursor control keys where you want to make insertions or deletions. When pressed hard, these keys will cause continuous movement of the cursor.

Errors detected after transmission can be corrected by re-typing the instruction.

2.5 THOSE LIGHTS ON THE SCREEN

There are eight indicator lights on the right hand side of the screen. They will go on and off as the screen and the computer communicate with each other - but you don't have to worry all that much about them. As long as you remember to press the ENTER key or one of the 12 special purpose keys to transmit your instructions and wait a few seconds for a response you should have no problems.

2.6 WHAT IF THE SYSTEM FAILS TO RESPOND?

If the computer fails to reply after one minute, press the RESET key, then press the ENTER key or one of the 12 special function keys. If you get no response after another minute, you should consult the project co-ordinator.

3. OPERATING THE TERMINAL

3.1 SIGNING-ON

The sign-on procedure is as follows:

* press the CLEAR key
* press the LOCK key
* type AQUA
* press the SHIFT key
* press the SPACE bar
* type PASS
* press the SPACE bar again
* type the two-character sign-on code assigned to you by the project co-ordinator
* adjust the brightness if necessary
* press the ENTER key

Signing-on is probably the only occasion on which you will need to remember to use both the LOCK and SHIFT keys. Once you have signed-on, your instructions and system responses will appear on the screen in upper case - regardless of whether you entered them in either upper or lower case. Instructions printed in upper case in this guide are printed so for the sake of clarity alone.

If you have followed this procedure correctly the system should respond with a display making up the word AQUARIUS - the name of STAIRS' on-line subsystem, then it will ask you which data base you wish to search. You can choose between RIE and CIJE, and in each case between one month's data and two years' data.

* Type R2YR,C2YR,RMTH or CMTH (the four options referred to above) followed by the ENTER key. STAIRS will then ask you which function you require

* type ..SET EXIT1 and press the ENTER key. This will invoke a system option set up for this implementation of STAIRS

* press the SEARCH, RANK, BROWSE or EXEC key to begin your search

3.2 CHANGING DATA BASES

You can select another data base to search by pressing the CHANGE key; but before you do this, please SAVE your queries, otherwise they will be purged from the system. We would like to review your final strategy to assess the feasibility of a decentralised information retrieval network in which users conduct their own searches.

3.3 SIGNING-OFF

You can sign-off at any time by typing ..OFF and pressing the ENTER key. If you sign-off with ..OFF CONT and if you sign-on again on the same day, your queries will still be available even if you forgot to SAVE them.

4. SEARCHING THE DATA BASE

4.1 STAIRS MODES AND FUNCTIONS

STAIRS has many moods: it can do many different things. The jargon term for these moods is "modes". The main modes are:

* HELP for seeking help
* SEARCH, RANK and SELECT for retrieving citations
* BROWSE for reading and printing retrieved citations

You can move from mode to mode simply by pressing the appropriate function key. When you return to one of the retrieved modes from another mode, your queries are still available. You only lose them if you CHANGE data bases or sign-off without using the SAVE function.

4.2 WHAT TO DO WHEN YOU NEED HELP

Press the HELP key. If you are not sure what to do next, by pressing the HELP key you will be asked to indicate what sort of help you require. For example, you may wish to learn more about a particular error message: by identifying the message to the system you will receive a more detailed explanation.

Over-use of the HELP function is fairly time-consuming and will reduce the effective time you have at the terminal. So please use HELP with some discretion and only when you are really stuck. Most times you will just have to press SEARCH or BROWSE to resume where you left off.

4.3 THE SEARCH MODE

You can get into the search mode by depressing the SEARCH key. STAIRS will then ask for your query, and at the same time assign it a number. These query numbers are very handy if you wish to refer back to earlier queries. The system stores the queries and you can ask for them to be displayed by pressing the DISPLAY key.

The simplest type of query is a single word such as CURRICULUM typed on the keyboard, followed by pressing the enter key. STAIRS will respond with a count of the number of times CURRICULUM occurs in the data base and will also post a tally of the number of citations in which CURRICULUM occurs.

4.4 SEARCH OPERATORS

This is more jargon. An operator defines the relationship between two or more terms for which you may be conducting a search. For example, by searching for CURRICULUM AND DEVELOPMENT you would retrieve all citations in which both the terms CURRICULUM and DEVELOPMENT occurred.

Common operators are:

OPERATOR	MEANING
AND	in the same document citation as
PARA	in the same paragraph (part of a citation) as
SENT	in the same sentence (part of a paragraph) as
ADJ	immediately followed (in the same sentence) by
OR	and/or
NOT	but not in the same document citation as

Please note that if you forget to specify an operator, the system will assume you mean OR. For example, if you type CURRICULUM DEVELOPMENT when you mean CURRICULUM ADJ DEVELOPMENT, STAIRS will interpret your entry as CURRICULUM OR DEVELOPMENT.

STAIRS will always display intermediate counts of term occurrences before displaying the final answer. Sometimes you will need to press the ENTER key to get the final result, and if the intermediate results fill more than one screen, you will need to press the ENTER key several times.

4.5 PARENTHESES

Since several words may be related by several operators, parentheses are often useful for retaining the sense of your search command. For example, CURRICULUM or CURRICULA ADJ DEVELOPMENT OR ENHANCEMENT would retrieve citations containing the phrases CURRICULUM DEVELOPMENT, CURRICULUM ENHANCEMENT, CURRICULA DEVELOPMENT or CURRICULA ENHANCEMENT. Parentheses will make this a little clearer: (CURRICULUM OR CURRICULA) ADJ (DEVELOPMENT OR ENHANCEMENT). You can see that this expression is quite different from: CURRICULUM OR (CURRICULA ADJ DEVELOPMENT) or ENHANCEMENT. Similarly, (CURRICULUM ADJ DEVELOPMENT) SENT CENTRE is much less ambiguous than CURRICULUM ADJ DEVELOPMENT SENT CENTRE.

As a rule-of-thumb, parentheses should be used in all search commands involving more than one operator.

4.6 BACK-REFERENCING

More jargon: this time to refer to the incorporation of an earlier query within a subsequent query. For example:

QUERY NO.	QUERY
1	CURRICULUM
2	DEVELOPMENT
3	1 ADJ 2
4	CENTRE
5	3 SENT 4

Query 5 refers to query 3 which refers to queries 1 and 2 (curriculum adjacent to development). That is two levels of back-referencing.

4.7 KNOWING WHERE YOU ARE

If you lose track of your query numbers, you can display your queries, complete with numbers and term occurrences, just by pressing the DISPLAY key. If there are too many queries to fit on one screen, you can view the next screen by pressing ENTER.

4.8 MASKING AND THE ROOT OPERATOR

Still more jargon. Instead of typing, say, CURRICULUM OR CURRICULA, you could type CURRICUL$. The '$' is called a masking character. CURRICUL$ will match any word beginning with the letters CURRICUL - including CURRICULA and CURRICULUM. Naturally, you need to be careful. READ$ will retrieve citations containing READ or READS or READING, but it will also retrieve citations containing READINESS or READILY or READY. Some control can be exercised by specifying the length of the mask. For example READ$3 will retrieve READILY but not READINESS. Please be careful with masking. With most of the words from two years of ERIC data at your fingertips, you will find that indiscriminate masking will yield many words with common stems. You may even exceed the capacity of the system in which case you will need to rephrase your query to limit the number of term occurrences.

A way out of this mess, and a way of making more efficient use of the system, is to use the special operator ROOT. Typing ROOT followed by a word root causes a display of all words beginning with just the root. For example, ROOT READ will yield a list of all words beginning with READ. Please note that the ROOT operator will only display word stems; it will not cause the retrieval of citations containing the specified word stems.

If you retrieve too many terms to be seen on one screen, just press the ENTER key as often as it is necessary to display successive screens.

4.9 STOPPED WORDS

When the ERIC files were inverted, certain words were suppressed, and no indexes were constructed on them. These were words with low information content such as a, at, and, the, of, etc. For this reason a query such as DEVELOPMENT ADJ OF ADJ CURRICULA would retrieve nothing. That is, there would be no term occurrences of 'of'. An alternative and potentially fruitful approach would be to search for DEVELOPMENT SENT CURRICULA.

In order to economise on the amount of direct access storage required for the Project, several hundred stop words have been identified. So it is possible that you will attempt to use at least one of them in a search. It will be evident from the displayed list of term occurrences that you have used a stop word, and you can confirm this by using the ROOT operator.

4.10 PARAGRAPHS

Document citations are made up of formatted fields (see Section 7.3) and unformatted text. The unformatted text is made up of paragraphs, sentences and words. The paragraphs are grouped into classes, and each class has a name:

	PARAGRAPH CLASS NAME	EXPLANATION
*	ACCN NO	Identification number
*	TITLE	Title
*	PUBN DATE	Publication date
*	CL HSE ACCN	Clearinghouse accession number
*	JNL CITN	Source journal title, volume, issue, pagination
*	ISSUE	Issue of RIE or CIJE

*	AUTHOR	Personal or corporate author
	INSTITUTION	Author affiliation
*	ABSTRACT	Narrative summary
	DESCR NOTE	Further descriptive information, series note, language, etc.
	SPONS AGENCY	Sponsoring agency
	AVAILABILITY	Source from which the full text of the document can be purchased
	PRICE	Whether the document exists in microfiche or hardcopy, price from the ERIC Document Reproduction Service
*	DESCRIPTORS	Terms from the ERIC Thesaurus. (See Section 10)
*	IDENTIFIERS	Additional subject terms not in the thesaurus. For example, project names

All these paragraph classes may occur in RIE document citations. The asterisk indicates the paragraph classes occurring in CIJE citations.

You may if you wish search particular paragraph classes. For example, SMITH.AUTHOR. will retrieve citations with SMITH in the author paragraph class. A double period may be used to specify negation, so that (CURRICULUM ADJ DEVELOPMENT).. DESCRIPTORS, IDENTIFIERS. will retrieve citations containing the phrase CURRICULUM DEVELOPMENT in a paragraph class other than DESCRIPTORS OR IDENTIFIERS.

5. WHAT TO DO AFTER THE SEARCH

5.1 LOOKING AT YOUR RESULTS

If you press the BROWSE key you can display retrieved citations on the screen. If you do this before you have submitted a search, you can browse through the whole data base. STAIRS will ask you how much of each citation you want to see. If you only want to browse through a few citations you can type ALL (followed by the ENTER key) and STAIRS will display all elements of the document citations. But if you have retrieved more than say, 5 citations, you probably will not want to view all the details of all the citations. In this case you can type C and STAIRS will display the formatted field and paragraph class names and numbers and ask you to indicate which parts of the citations you wish to see. For example, if you type 2 the display will contain only titles. If you type 1,2,3, accession numbers, titles and publication dates will all be displayed.

5.2 MOVING AROUND YOUR ANSWER SET

Once you have entered the browse mode you can view your results in just about any order you specify. Pressing the ENTER key will cause the next screen to be displayed. The next screen is also called the next page. If you entered P=5, STAIRS will skip to and display the fifth page of your results. If when you had viewed the fifth page you wanted to move back to your starting point, you have only to enter P=1. Similarly with document citations. DOC=10 means "skip to document citation number 10". DOC=5 means skip back to document citation number 5, and so on.

5.3 PRINTING YOUR RESULTS

Browsing not only enables you to identify useful terms for subsequent searches, it also helps you choose the retrieved citations you wish to print. You have the option of printing your results at the terminal or at the central computer installation in Canberra.

Let us assume that you have decided which paragraphs you want to print, and that you want to print the citation appearing on the screen. All you do is press the COPY key. If you want to print, say, three citations at the terminal you will need to display them one at a time on the screen and press the COPY key once for each citation you want printed. Since the printing operation only takes place after you have signed-off, you will need to allow time for this at approximately two minutes per citation within your scheduled terminal session. For example, if you are scheduled for one hour on the system and you have asked for 10 citations to be printed at the terminal, you will need to sign-off no later than 40 minutes into your session.

If on the other hand you wish to use all your scheduled time for interaction with the system, you will find the off-line print on the high speed printer in Canberra a very useful facility. Say you wish to print all the document citations in your answer set: by entering ..PRINT ALL, the computer will print these in Canberra. They will then be mailed as soon as possible to the address on your questionnaire. You can also use the off-line printer for selective printing of your results. For example, say you had completed a search, browsed through the results, and decided that you wanted to print citations 2,3,4 and 17. You will need to issue two separate print commands: ..PRINT 2,4 to print citations 2,3 and 4; and .. PRINT 17 to print the seventeenth citation.

Please remember that if you want to order the full text of the document for which you have retrieved the citation, you will need to respecify your print requirements to include accession numbers. These numbers uniquely identify each document and must be quoted in your order.

6. DOCUMENT BACK-UP

6.1 GETTING HOLD OF ERIC DOCUMENTS

Most reports listed in Resources in education are available in microfiche format from the National Library. Most journal articles indexed in Current index to journals in education are available as photocopies from the National Library.

Ordering microfiche - RIE items

All the reports listed in RIE have unique ED (ERIC Document) numbers and orders for ED items should include the ED number and report title. You may order them as duplicate microfiche (at 20 cents per fiche) or as photocopies made from microfiche (at 6 cents per exposure). The Library can provide a faster service if requests are arranged in ascending order of ED numbers. An invoice will accompany each order despatched.

Please label your order ON-LINE ERIC PROJECT and send it to:

Education Information Service
National Library of Australia
Canberra ACT 2600

Telephone: (062)621 111 Telex: LIBAUST 62100

Ordering photocopies - CIJE items

All the journal articles indexed in CIJE have unique EJ (ERIC Journal) numbers and orders for EJ items should include the EJ number, journal title and issue, and article title and author. To further complicate things, they should be made in triplicate with only one item per page. Each completed order will be despatched with an invoice and you will be charged at the rate of 5 cents per photocopied page.

Please note that the supply of photocopies will be subject to the provisions of the Copyright Act.

Please label your order ON-LINE ERIC PROJECT and send it to:

Loans and Locations Service
National Library of Australia
Canberra ACT 2600

Telephone: (062)621483 Telex: LIBAUST 621000

You might find that your requirements for photocopied journal articles can be met by your own library. If this is the case, we would prefer that you tried your local sources first, falling back on the National Library only when these sources lack depth or coverage.

7. OTHER FEATURES OF STAIRS

7.1 EXPLANATION

The description of STAIRS so far has concentrated upon a basic set of instructions - sufficient for you to search, review and print citations from the ERIC data base. STAIRS however is quite a sophisticated system and some of its additional functions are described below. Your project co-ordinator has detailed manuals should you wish to really explore the system.

7.2 RANKING YOUR RESULTS

The RANK key can be used like the SEARCH key except that it sequences retrieved items so that in BROWSE mode the most relevant items appear first. Relevance is determined by STAIRS on the basis of the number of query terms present in the retrieved citation. You have a choice of five different ranking algorithms, each one of which calculates relevance in a slightly different way.

7.3 SELECTING ITEMS

Whereas the SEARCH and RANK functions retrieve citations by identifying a match between query terms and words in the data base, the SELECT function retrieves citations by virtue of the contents of certain formatted fields.

These fields are:

FIELD NAME	EXPLANATION
DATE	Year of publication (2 digits)
PAGES	Number of pages in whole document
TYPE	Publication type code (see below)

Only DATE is available for CIJE items, and TYPE is not available for RIE items before September 1974 (it wasn't included until then). Where the number of pages is unspecified and consequently unknown, 999 is assumed as a default value.

Publication type codes

Code	Meaning
A	Non-print media, eg. AV materials
B	Books
C	Curriculum guides, teacher-developed materials, laboratory manuals
D	Directories, membership lists, reference works dealing with organisations and institutions
G	Guides: teaching guides, resource guides, study guides, training manuals
H	Legislation - including hearings and reports
J	Journal articles, serials, periodicals, bulletins, newspapers
K	Project descriptions
L	Bibliographies, catalogues, state-of-the-art reviews, book reviews, indexes, literature guides, abstracts
M	Maps, atlases, gazetteers
N	Numerical and statistical tables, quantitative data and analyses
O	Other
P	Proceedings, conference records/minutes
Q	Questionnaires, tests, measurement and evaluation devices
R	Reports (research and technical)
S	Speeches and other verbal presentations
T	Theses, dissertations
V	Dictionaries, vocabularies, glossaries, thesauri
Y	Annual reports, yearbooks

The set of document citations retrieved for a particular query (the raw material of the SELECT function) is called an answer set. When you press the SELECT key STAIRS will ask for the number of the answer set from which the selection is to be made, and will ask for the SELECT parameters. For example, the instruction 3 TYPE EQ S (followed by the ENTER key) is a request for citations to be selected from the answer set of query number 3 where they have type "S" ("speeches and other verbal presentations").

The SELECT function has its own operators:

OPERATOR	MEANING
EQ	Equal to
NE	Not equal to
GT	greater than (Z is "greater" than A)
LT	less than
NL	greater than or equal to
NG	less than or equal to
WL	within the limits (see below)
OL	outside the limits (see below)

WL and OL require two values separated by a comma. For example, if you typed 5 TYPE WL P,T (then ENTER), citations would be selected from the answer set of query 5 if their TYPE was P,Q,R,S, or T.

8. SAVING QUERIES

You may wish to save queries to use on a different data base or at a later date. The way you do this is to:

- DISPLAY your queries
- remove unwanted queries using the ..PURGE command. For example: ..PURGE 2 will remove your second query ..PURGE 3,5 will remove queries 3,4 and 5. Please remember not to purge queries which are referred to by queries you wish to retain. For example: say you wanted to save query number 15 and query 15 referred to queries 13 and 14. In this case you would need to save queries 13,14 and 15.
- SAVE wanted queries
- CHANGE data bases, or sign-off (and sign-on again at your next scheduled terminal session)
- EXEC (execute) the saved queries

To SAVE your wanted queries you will need to enter

..SAVE

STAIRS will then ask for a four character name under which the queries will be stored. The project co-ordinator will allocate a name to you. You will need to remember it to retrieve your queries.

9. SEARCHING SAVED QUERIES

When you press the EXEC key, STAIRS will ask for the four character name. Type it in, press ENTER and your stored queries will be searched. To search your next stored query, just press the ENTER key again.

10. THE ERIC THESAURUS: a guide to its contents and use

WHAT IS IT?

The Thesaurus of ERIC Descriptors is an alphabetical listing of terms used by ERIC indexers to describe the subject contents of the documents they index. It contains approximately 8000 descriptors.

WHY YOU SHOULD USE IT

Although it is not imperative that you thoroughly familiarise yourself with the contents of the Thesaurus, it is important that you are aware of its structure and applications. There is no doubt that the Thesaurus is the single most useful aid for searching the data base. By referring to its vocabulary which is carefully controlled you can ensure that the descriptors you choose to express your query are the same as the descriptors the ERIC subject specialists have used to describe your subject. Your efforts to learn this language will be handsomely repaid in your results.

However, if you are interested in very new or a highly specific topic you may find that there are no Thesaurus terms which accurately describe them. In such cases, STAIRS can really prove its worth. With STAIRS you always have the option of searching for any terms anywhere in the document citation. This is called a free-text capability. You do not always have to search for Thesaurus descriptors.

LAYOUT OF THESAURUS ENTRIES

In the main sequence of the Thesaurus, the descriptors appear in capital letters, filed alphabetically, letter by letter. The entry for each descriptor defines the hierarchical and lateral relationships between the entry descriptor and other descriptors in the vocabulary.

NOTATION OF THE MAIN SEQUENCE

These term-term relationships are expressed in a notation common to most thesauri. Six symbols are used: USE; UF; NT; BT; RT and SN.

(i) USE

Example: TROUBADOURS
USE Poets

The "USE" symbol directs you from a rejected term to an accepted term (poets). In other words, you are cautioned that "troubadours" is not a word in the ERIC language, and advised to use "poets" instead.

(ii) UF

Example: POETS
UF Troubadours

This is the converse of the "USE" symbol. It merely informs whoever is using the Thesaurus that the descriptor "poets" is to be used for, or instead of, the descriptor "troubadours".

(iii) NT

Example: LEAVE OF ABSENCE
NT Sabbatical Leaves

"NT" means "narrower term". It is used to indicate specific instances of general descriptors. It denotes a hierarchical relationship and says in effect, "if you are interested in 'leave of absence' - you may also be interested in a specific instance, namely 'sabbatical leaves'".

(iv) BT

Example: SABBATICAL LEAVES
BT Leave of Absence

"BT" means "broader term". It too denotes a hierarchical relationship, but this time the direction is from the particular to the general. This is the complement to the "NT" symbol and says in effect, "if you are interested in 'sabbatical leaves' you may also be interested to learn that 'leave of absence' is an alternative broader term".

(v) RT

Example: LASERS
RT Electronics

"RT" means "related term". It denotes a lateral relationship between the main descriptor and other potentially useful descriptors and provides the context in which to interpret the main descriptor.

(vi) SN

Example: LEARNING MODALITIES
SN Refers to the sense modality used by the learner when instruction is presented in different sensory stimulus modes.

"SN" means "scope note". Scope notes explain the denotation and connotation of potentially ambiguous descriptors.

OTHER SECTIONS IN THE THESAURUS

(1) The Rotated Descriptor Display

This is a KWIC (Key Word In Context) index to descriptors in the main sequence. Since it tends to group descriptors generically it can be very useful for determining whether your queries are exhaustive in their term specification. For example the entry:

```
        ADMISSION (SCHOOL)
COLLEGE ADMISSION
        ADMISSION CRITERIA
  EARLY ADMISSION
        ADMISSIONS COUNSELORS
```

displays the "admission terms" continuously. You don't have to search through the alphabetical sequence and its reference structure to discover possible omissions from your search strategy.

(2) Descriptor Group Categories

There are 52 descriptor groups covering such term classes as behaviour, curriculum, languages, psychology and standards. The 52 descriptor group categories function more or less as scope notes for the descriptors themselves. For example:

160 Environment

Aggregate of conditions or influences on communities, schools, culture, and social factors, eg. Community Influence, Classroom Environment, etc. See also CULTURE, SOCIOLOGY.

The Descriptor Group Categories really only make sense when considered with the

(3) Descriptor Group Display

Each descriptor group is a list of subject-related descriptors. For example:

160 Environment

ACOUSTICAL ENVIRONMENT
CITY IMPROVEMENT
CLASSROOM ARRANGEMENT
CLASSROOM ENVIRONMENT
CLIMATE CONTROL
CLIMATIC FACTORS etc.

This can be a useful device for demonstrating relationships between descriptors. It is, if you like, a kind of classified index to the Thesaurus.

USING THE THESAURUS

Let us assume for the sake of this example that you are interested in the evaluation of secondary school curricula. What to do next?

1. Pick the words and phrases that define the subject.

   evaluation
   curricula
   secondary schools

2. Check in the Thesaurus to ensure that you are using the same language the ERIC indexers used.

Query terms		Thesaurus terms
evaluation	stays	evaluation
curricula	becomes	curriculum
secondary school	becomes	secondary schools

3. Next choose Thesaurus descriptors to further delimit your topic. Consider the addition of broader, narrower and related terms to your query. Under the entry for evaluation are two narrower terms: course evaluation and curriculum evaluation. Similarly under secondary schools is the narrower descriptor "high schools".

By assembling these terms in a logical order we arrive at a query statement which looks promising:

(COURSE EVALUATION OR CURRICULUM EVALUATION)
AND
(SECONDARY SCHOOLS OR HIGH SCHOOLS)

All that remains is to search the query against the data base. Its effectiveness can only be judged in the light of the citations it retrieves.